Novel plant DNA binding protein: Non-expresser of pathogenesis related 1 gene (NPR1) involved in disease resistance

Saba Pirnia, M.Sc.

A Thesis

submitted to the Center for Biotechnology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Brock University

St. Catharines, ON

ABSTRACT

For the first time, through validation of the modified Chromatin Immunoprecipitation (ChIP) method (in vitro ChIP), the direct binding of NPR1 to the PR1 promoter was demonstrated. This is a novel advancement on plant systemic acquired resistance (SAR)-mediated disease responses. The NPR1 protein (non-expressor of pathogenesis related gene 1) is a transcriptional co-activator and positive regulator of SAR, a long-lasting mobile defense signal found in plants. The pathogenesis-Related gene 1 (*PR1*) is particularly induced during defense response, and as such, is typically used as a marker for establishment of SAR in plants. Salicylic acid (SA) is a phytohormone required for SAR-mediated defense responses against pathogens. Recently, the role of NPR1 as a SA receptor was demonstrated; SA has been shown to directly bind to NPR1 through Cysteine 521 and 529 on the C-terminus region of NPR1 via the transition metal copper. The binding of SA to NPR1 results in disruption of the interaction between BTB/POZ and the C-terminus domains of NPR1. Upon SA-NPR1 binding, the C-terminus transactivation domain is released from the auto-inhibitory BTB/POZ domain, resulting in activation of the NPR1 transcription co-activator function, followed by *PR1* transcription in *Arabidopsis thaliana*. *Arabidopsis thaliana* has an inducible defense system and is considered a model plant for studying disease resistance responses.

In the current research, NPR1 was demonstrated to bind to the PR1 promoter at two distinct regions, in the presence and absence of SA. In the presence of SA, the binding site of NPR1 was determined to be localized at the -636 to -646 base pair sequence; however, in the absence of SA, NPR1 was found to bind around the -790 to -833 base pair sequence. In addition, two distinct DNA binding domains were identified within NPR1, localized on the C and N-terminus regions. In the absence of

SA, the DNA binding domain within the N-terminus region, located between amino acids 110-190, was shown to facilitate the binding of NPR1 to the PR1 promoter through the amino acid cysteine 150 (Cys^{150}) via transition metal. The DNA binding domain on the C-terminus region, located between amino acids 513-535, was demonstrated to allow the binding of NPR1 to the promoter of PR1 in the presence of SA. Two amino acids, cysteine 521 and 529 ($Cys^{521/529}$), were shown to be essential for SA binding to NPR1 and subsequent NPR1 binding to the PR1 promoter. Furthermore, 4hydroxy benzoic acid (4-OH-BA), the inactive analogue of SA, has been demonstrated to be a potent inhibitor of NPR1-PR1 promoter interaction, both in vivo and in vitro, by competing with SA for NPR1 binding. Moreover, we demonstrated that other analogues of the NPR1 protein, NPR2, NPR3, and NPR4, are also recruited to the PR1 promoter. NPR4 showed a similar binding profile to NPR1, both in the presence and absence of SA. NPR2 and NPR3 were observed to only interact with the PR1 promoter in the absence of SA. Both NPR5 and NPR6 were shown to forgo binding to the PR1 promoter, further confirming their role in plant developmental processes other than defense. In addition, the binding of NPR1 to the PR1 promoter was demonstrated to be conserved among other plant species, including rice and maize. Both rice and maize NPR1 proteins were observed to bind to the PR1 promoter in the presence of SA and a metal co-factor, similar to *Arabidopsis* NPR1.

Our results expand our understanding of how NPR1 interacts with the PR1 promoter to regulate gene expression during SAR establishment. This study also revealed that NPR1-mediated SAR defense signaling is conserved among other crop species, which can potentially facilitate the identification of novel plant-priming compounds through high-throughput chemical screening procedures alongside the application of the validated in vitro ChIP technique as a primary screening method.

ACKNOWLEDGEMENTS

It is hard to overstate my gratitude to my supervisor, Professor Alan Castle. Without his inspirational guidance, his enthusiasm, his encouragements, his unselfish help, his endless support I could never finish my doctoral work in Brock University. For me, he is not only a teacher, but also a lifetime advisor.

I also want to thank my committee member, Prof. Travis Dudding who has always provided thoughtful comments and suggestions.

Special thanks to Professor Jeffrey Atkinson, who believed in me when I started my PhD program at Brock University. I will forever be indebted to Dr. Atkinson for investing in my professional and personal development and for his inexhaustible support, encouragement, enthusiasm, and patience in leading me to achieving my goal. Throughout the years that it took to complete my thesis I could always lean on Dr. Atkinson for wisdom and expertise to keep me on the right path.

My special thank goes to my external examiner Dr. Keiko Yoshioka for her kindness and patience in going through my PhD defence and for her brilliant comments and suggestions.

I would like to express my special appreciation and thanks to Director of Graduate Studies, Ms. Gail Pepper. I would like to thank you for encouraging my research and for your patience, motivation, and enthusiasm.

Additionally, I want thank all the members of Dr. Yan's lab since beginning grad school for the rich experiences we've had together. Special thanks to Ravi Shekar for being a great friend who brought up engaging conversations on variety of chemistry/biology topics and memorable friendship.

I gratefully acknowledge the funding sources that made my Ph.D. work possible particularly, Dean of Graduate Studies Research Fellowship.

Finally I would like to thank my mother Parvin, my dad Reza, my sisters, Sima and Sahar, and my brother Amir for their unconditional love and support through all these years. I am proud to be raised in a family, appraising higher education. Thank you for all the sacrifices you made, the life principles you taught me, and the comfort you provided me. Without you and your support I could not have achieved this.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 3

APPENDIX

LIST OF FIGURES

CHAPTER 2

CHAPTER 4

CHAPTER 5

APPENDIX

LIST OF ABBREVIATIONS

°C - degrees centigrade

% - Percentage

μ - micro-

4-OH-BA - 4-hydroxy benzoic acid

4AzSA - 4-AzidoSA

3-OH-BA - 3-hydroxy benzoic acid

AA - amino acid

ABA - abscisic acid

AAO- Arabidopsis aldehyde oxidase

AOBP- Ascorbate Oxidase gene-binding protein

ARD - ankyrin repeat domain

Arg - arginine

as-1- activation sequence 1

Asn - asparagine

Avr – avirulence

AzA- Azelaic acid

AZF1- Arabidopsis zinc finger protein 1

BA- benzoic acid

BA2H - benzoic acid 2-hydroxylase

BL- brassinosteroid

bHLH - basic helix-loop-helix

BOP1- BLADE-ON-PETIOLE1

BOP2- BLADE-ON-PETIOLE2

BTB/POZ - Broad-Complex, Tramtrack, and Bric-a-brac/Pox virus and Zinc finger

BTH - benzo-l ,2,3-thiadiazole-7 -carbothioic acid S-methyl ester or benzothiadiazole

bZIP - basic domain/leucine zipper

CaMV - cauliflower mosaic virus

CAP- catabolite activator protein

ChIP - chromatin immunoprecipitation

CPB60g- calmodulin-binding protein 60g

CPL- chorismate pyruvate-lyase

CT C-terminus

CK- cytokinin

Cys – cysteine

D - aspartate

DB - DNA binding domain

DA- dehydroabietinal

DIRl - defective in induced resistance 1

DNA deoxyribonucleic acid

DOF- DNA-binding with one finger

DSC- Differential scanning calorimetry

dsDNA - double-stranded DNA

DTT – dithiothrietol

E - glutamate

E. coli - Escherichia coli

EDTA - ethylenediamine-tetraacetic acid

EMSA - electromobility shift assay

ET – ethylene

ETI - effector-triggered immunity

FPLC - fast protein liquid chromatography

g - gram(s)

G3P glycerol-3-phosphate

GAL – galactosidase

GA- gibberellin

Gly - Glycine

GSNO - S-nitrosoglutathione

GSSG - oxidized glutathione

GUS- β-glucuronidase

His - Histidine

HR - hypersensitive response

hr - hour(s)

HTH- Helix-turn-helix

ICS - isochorismate synthase

IEGTl - immediate-early-induced glucosyl transferase 1

IHF- integration host factor

INA - 2,6-dichloroisonicotini acid

IP - immunoprecipitation

IPL - isochorismate pyruvate lyase

ISR - induced disease resistance

IRS- sequences with inverted repeats

ITC- isothermal titration calorimetry

JA - jasmonic acid

Km - kanamycin

l - litre(s)

Leu – leucine

Lr10- leaf rust resistance 10

LS - linker scanning

LS5- linker scanning 5

LS7- linker scanning 7

M - molar

min - minute(s)

mol - moles

MAP - mitogen-activated protein

MAMPs - microbe-associated molecular patterns

MeJA - methyl jasmonate

MeSA - methyl salicylate

nahG - salicylate hydroxylase

NDR1- Non-race-specific Disease Resistance 1

NF-$_K$B- nuclear factor kappa-light-chain-enhancer of activated B cells

NLS - nuclear localizing signal

NPR1 - non-expressor of pathogenesis related genes 1

NT N-terminus

P- proline

PAD4- Phytoalexin Deficient 4

PAGE - polyacrylamide gel electrophoresis

PAL - phenylalanine ammonia lyase

PCR - polymerase chain reaction

PDF1.2 - plant defensin 1.2

PGRP- plant growth-promoting rhizobacteria

Phe- phenylalanine

PMA- protein microarray

PR - Pathogenesis related

PRRs - pattern recognition receptors

PVDF-Polyvinylidene fluoride

qPCR - quantitative PCR

R – resistance

RAV1- Related to ABI3/VP1

RAV2- Related to ABI3/VP2

rbcS- ribulose bisphosphate carboxylase

rcf - relative centrifugal force

RLK - receptor-like kinases

RNAi - RNA interference

RNA- ribonucleic acid

ROS - reactive oxygen species

Rpg1- Ruptured pollen grain 1

rpm - revolutions per minute

RT- room temperature

S - serine

SA - salicylic acid

SAH- salicylate hydroxylase enzyme

SAMT1 - SA-Methyl transferase1

SAP- stress-associated protein

SAR - systemic acquired resistance

SARD1 - SAR Deficient 1

SDS - sodium dodecyl sulfate

SD - standard deviation

sfd1- Suppressor of fatty acid desaturase deficiency 1

SGE SA glucose ester

SN- Supernatant

SREBP - element-binding protein 1

ssDNA - single-stranded DNA

TA - transactivation domain

TATA-TATA box

TGA - TGACG motif binding

Tyr - tyrosine

t-CA - trans cinnamic acid

TE - Tris buffer and EDTA

TF - transcription factor

TMV tobacco mosaic virus

TRF1- telomeric repeat binding factor 1

Trp – Tryptophan

UBQ- ubiquitin

Why 1 - Whirly 1

WRKY- WRKYGQR motif binding

WT-wild type

Y-tyrosine

ZAT10- zinc transport of Arabidopsis 10

ZF - zinc finger

CHAPTER 1- INTRODUCTION

Salicylic acid (SA) is an endogenous plant phytohormone mandatory for establishment of systemic acquired resistance (SAR), a broad-spectrum and long-lasting plant immune response that occurs following exposure to pathogens (Cao et al., 1994; Vlot et al., 2009). In plant immune response, the NPR1 (Non-expressor of pathogenesis related gene1) protein is responsible for positive regulation of SAR-mediated immune responses. The *PR1* gene is a molecular marker for SAR; in unchallenged cells, its transcription is repressed by the TGA2 transcription factor (Zhang et al., 2003; Rochon et al., 2006). However, upon pathogen attack and subsequent SA accumulation, NPR1 act as a transcriptional co-activator, activating *PR1* expression through the formation of an enhancesome with TGA2 on the promoter of PR1 and masking TGA2 repressor activity (Boyle et al., 2009; Rochon et al., 2006). Consequently, plants lacking functional NPR1 are unable to express the *PR1* gene and activate SAR; as such, they are more susceptible to pathogen attack (Cao et al., 1994; Delaney et al., 1995).

Recently, two independent groups demonstrated that NPR1 is the receptor for SA (Wu et al., 2012; Manohar et al., 2015). SA binds NPR1 directly through two cysteines (Cys521 and Cys529) located in the C-terminus region (Wu et al., 2012) via the transitional metal copper. Notably, NPR1 is the first example of a plant transcription regulator that also binds copper. Binding of SA to NPR1 causes a conformational change in NPR1 and disrupts the interaction between the BTB/POZ and transactivation domains; the C-terminus transactivation domain is released from the BTB/POZ auto-inhibitory domain, enabling NPR1 to activate *PR1* transcription on the PR1 promoter (Wu et al., 2012).

Various research papers have reported that SA-dependent, NPR1-mediated defense responses are conserved among multiple plant species, including tomato (Lin et al., 2004), wheat (Makandar et al., 2006), rice (Chern et al., 2005; Yuan et al., 2007), soybean (Sandhu et al., 2009), banana (Endah et al., 2008), cotton (Zhang et al., 2008), and tobacco (Maier et al., 2010). However, the cysteine residues Cys521 and Cys529, required for the binding of SA to *Arabidopsis* NPR1, have been shown to not be conserved in NPR1 proteins from other crops (Wu et al., 2012).

Recently, additional members of the NPR protein family, including NPR3 and NPR4, were also shown to interact with SA and negatively regulate SAR (Liu et al., 2005; Zhang et al., 2006; Fu et al., 2012). Indeed, Fu and co-workers (2012) postulated that NPR3 and NPR4 act as CUL3 E3-ligase substrate adaptors and recruit NPR1 for degradation by proteasome. However, despite the remarkable insights on NPR3 and NPR4 functions provided by different groups, the inconsistencies in pathology around the *npr3* and *npr4* mutant plants have cast doubts as to the reliability of these findings, as discrepancies in results have surfaced between reports submitted by different groups of researchers. For example, in research conducted by Fu et al. (2012), disease resistance results obtained for single *npr3* or *npr4* mutant plants were shown to not diverge significantly from those obtained for wild-type plants. Similarly, Zhang et al. (2006) demonstrated that *npr4-3* and *npr4-2* plants were equally resistant to bacterial pathogens as compared to wild-type plants. Contrastingly, Liu et al. (2005) reported that disease resistance in *npr4-1* plants was compromised upon pathogen challenge. Chiefly, the redundant conclusion from these groups is that both NPR3 and NPR4 are associated with plant immunity.

1.1 Objectives

The goal of the currently presented research is to provide a better understanding of the role of NPR1 in SAR-mediated immune responses, specifically focusing on its interaction with the PR1 promoter, and the behavior of this complex. This research has broad implications, as the findings from the study of *Arabidopsis*, a model plant, can be applied towards other crops, including rice and maize.

Although preliminary results from Rochon et al. (2006) indicated that NPR1 is recruited to the PR1 promoter *in vivo,* the processes involved in the direct interaction between NPR1 and the PR1 promoter remained to be elucidated. Thus, the validation and subsequent application of an appropriate methodology could provide answers regarding the binding domains of NPR1, allow for localization of the NPR1-PR1 promoter binding sites, and clarify the role of SA and its inactive analog 4-OH-BA in these interactions. Furthermore, such a method could be used to determine whether NPR1 from different crops also binds SA through metal coordination, similar to *Arabidopsis* NPR1. Lastly, application of a reliable method could help define the role of other NPR protein family members, including NPR3 and NPR4; particularly, it could help answer questions in regards to their binding preferences, shed light on the relationship between NPR1, NPR3, and NPR4 with respect to disease resistance and SA interaction, and clarify inconsistencies in results obtained by different groups.

1.2 Outline

Chapter 2 offers a comprehensive literature review that provides general background information regarding plant immunity, focusing on systemic acquired resistance (SAR)-mediated non-expressor of pathogenesis related gene 1(NPR1) and salicylic acid (SA) disease resistance responses. The second half of the literature review emphasizes the principles behind protein and DNA interactions, including recognition mechanisms, DNA topology, and the complexities involved in such interaction processes. Additionally, examples of different DNA binding protein families, as well as methods for studying protein DNA interactions, are discussed in detail.

Chapter 3 provides a detailed description of all methods and techniques used to generate the data presented in chapter 4.

Chapter 4 includes the most significant results on the binding of NPR1 to the PR1 promoter. Section 4.1 explains the validation of the modified *in vitro* Chromatin Immunoprecipitation (ChIP) method used to investigate the binding of NPR1 to the PR1 promoter. Sections 4.2 to 4.5 illustrate the binding sites of NPR1 on the PR1 promoter in the presence of SA, as well as the effect of other SA analogs, including 4-OH-BA and INA, on the binding interactions of NPR1 with the PR1 promoter. Sections 4.6 to 4.8 focus on determinations of DNA binding domain locations on NPR1 in the absence of SA, and metal transition coordination. Section 4.9 emphasizes the specific characteristics of DNA required for NPR1 binding. Sections 4.3 and 4.10 expound on the relationship between 4-OH-BA and SA, as well as provide *in vitro* and *in vivo* observations of its effect on the binding of NPR1 to the PR1 promoter. Section 4.11 provides evidence that other NPR protein family members also bind to the PR1 promoter. This section particularly focuses on the

importance of NPR3 and NPR4 in SAR-mediated immune responses. The most relevant findings regarding the role of the NPR1 protein in other crops, including rice and maize, are presented in the last section (4.12).

Chapter 5 provides a detailed discussion on the validation of the new modified ChIP method, including comparisons with other available methods for studying protein DNA interactions. The newly established model, which illustrates the interaction sites on the PR1 promoter and the DNA binding domain locations of NPR1 with respect to SA, is explained and compared with relevant literature discussed in chapter 2. Moreover, contradictory findings from the literature on the roles of NPR3 and NPR4 in SAR establishment, as well as their interaction with SA, are discussed and compared to the currently presented findings. Lastly, the role of the NPR1 protein in other crops with respect to SA and metal co-factors is explained and compared with relevant literature discussed in chapter 2.

Chapter 6 summarizes all major findings and suggests future investigations needed to further advance our current understanding of NPR protein functions in SAR establishment.

CHAPTER 2- LITERATURE REVIEW

2.1 The plant innate immunity

In their natural environment, plants are constantly exposed to a variety of microorganisms, many of which are pathogens. Unlike animals, plants lack specialized immune cells; when a pathogen attack occurs, the surveillance systems of plants act by recognizing conserved microbial structures or effectors that are released from pathogens. For example, the physical barriers of plants, such as waxy cuticles or lignified cell walls, secrete secondary metabolites that repel many potential pathogens (Fan and Doerner, 2012).

This surveillance system, in charge of detecting the presence and nature of infections, and the plants' first line of defense against pathogens, is referred to as 'plant innate immunity' (Medzhitov, 2001). The surveillance systems within plants are present in the form of receptors, which allow for a suitable and rapid host-defense response to be triggered upon attack (Ryan et al., 2007). When this system is successfully activated in a given plant, it can limit the growth of pathogens (i.e., increased plant resistance), resulting in what is referred to as an incompatible interaction.

To counteract this response, some pathogens release specific effectors capable of disrupting the plant's recognition pathways. A failure in recognition may lead to inactivation or a delay in the activation of appropriate defense responses, thus making the plant susceptible to infection, a process referred to as a compatible interaction (Jones and Dangl, 2006; Rasmussen et al., 2012). Generally, plants recognize pathogens through two different layers of receptors: resistance (R) proteins and pattern recognition receptors (PRRs) (Figure 1). Pattern recognition receptors, such as

transmembrane receptor-like kinases (RLK) and transmembrane receptor-like proteins (RLP), form the first layer of recognition. In the *Arabidopsis* gene family, around 600 genes encode for RLK, and 57 genes encode for RLP (Segonzac and Zipfel, 2011). These receptors recognize the conserved structures of pathogens, collectively resulting in the activation of pathogen- or microbe-associated molecular patterns (PAMPs or MAMPs) (Segonzac and Zipfel, 2011). Plants PPRs respond to MAMPs or PAMPs and generate PAMP-triggered immunity (PTI) (Figure 1), the process responsible for protecting plants from a wide range of pathogens (Nurnberger and Kemmerling, 2006). The second layer of recognition is comprised of *R*-gene encoded receptors that recognize effectors secreted from pathogens (Figure 1). Such effectors target different plant proteins associated with defense signaling pathways that lead to PTI suppression. At the same time, R proteins, which correspond to these effectors, activate effector-triggered immunity (ETI), generating a higher degree of defense signals (Figure 1) (Jones and Dangl, 2006). Effector proteins can develop escape strategies that allow them to go undetected by R proteins; as such, it is very important for R proteins to maintain their resistance role against pathogens. Any failure in keeping such balance may lead to an increase in the susceptibility of the plant against pathogens.

2.1.1 PAMP-Triggered Immunity (PTI)

At the local infection site of plants, there are multiple pattern recognition receptors (PRRs) tasked with recognizing PAMPs (Zipfel et al., 2006), such as flg22 (Gómez-Gómez and Boller, 2002), cell wall fragments chitin and beta-1, 3-glucan oligomers, and elongation factor Tu (Boller and Felix, 2009; Zipfel et al., 2006). Once the plant's signaling pathways are activated through PRRs, PTI-associated defense

responses are triggered (Jones and Dangl, 2006) (Figure 1). Upon pathogen encounter and activation of plant defense response, FLS2 heteromerizes with LRR receptor kinase BAK1 (BRI1 associated receptor kinase 1), followed by phosphorylation of both FLS2 and BAK1, an essential step in transducing an appropriate defense response (Gomez-Gomez et al., 2001 and 2002). The FLS2-BAK1 signaling response triggers a MAP kinase cascade (Chinchilla et al., 2007), MAPKKK - MKK4/MKK5 - MPK3/MPK6, leading to the activation of transcription factors such as WRKY22 and WRKY29, which in turn results in increased resistance against bacterial infection.

In potato plants, the PAMP Pep-13 is recognized by PRR. Pep-13 are highly conserved 13-amino acid fragments localized in the cell walls of the *Phythophthora* species (Brunner et al., 2002). Upon detection of Pep-13, PTI is triggered, resulting in a series of downstream defense responses that include H_2O_2 accumulation, activation of both salicylic acid (SA) and jasmonic acid (JA) hormones, and expression of related defense genes (Halim et al., 2009). In *Arabidopsis*, the chitin elicitor-binding protein recognizes an element in the pathogen cell wall named chitosan; this triggers a MAP kinase cascade, and subsequently, the activation of defense signaling responses, such as the generation of reactive oxygen species (ROS), Ca^{2+} and ion fluxes, callose deposition, and the production of the plant hormones jasmonic acid (JA) and abscisic acid (ABA) (Iriti and Faoro, 2009; Nicaise et al., 2009; Tsuda et al., 2008).

For successful pathogen resistance to occur, defense responses are coupled with the expression of related genes, the production of antimicrobial effectors, and the activation of the hypersensitive response (HR) (Pieterse et al., 2009). HR response causes rapid cell death at the local infection site, which leads to pathogen growth inhibition (Mur et al., 2008). Furthermore, hydrolase and hydrolase inhibitors are involved in PAMP-triggered immunity. After a pathogen attack occurs, during

recognition events, plants release hydrolases, including beta-1, 3-glucanase, and chitinase, to degrade the cell wall components of the pathogen, which subsequently activates PTI. At the same time, pathogens secrete different enzymes to degrade plant cell walls, such as polygalacturonase and pectate lyase. In response, plants release hydrolase inhibitors to degrade the enzymes that were secreted by the pathogens (Lagaert et al., 2009).

2.1.2 Effector-Triggered Immunity (ETI)

Pathogen effector proteins are able to suppress PTI and enhance the infection processes through adhesion to plant cells, degradation of plant physical barriers, and pathogen dispersion. Effector-triggered immunity (ETI) is a much more comprehensive process than PTI defense signaling, and activated by pathogen effectors called avirulence (Avr) proteins that are recognized by plant R proteins (Chisholm et al., 2006; De Young and Innes, 2006; Jones and Dangl, 2006) (Figure 1). R proteins are specific receptors similar to PRRs that detect general components of pathogens. Activation of ETI can also result in a hypersensitive response. Similar to PTI, after recognition events are triggered between plant R proteins and pathogen effectors, rapid production of reactive oxygen species and an increase in levels of Ca^{2+} concentration occur in host plant cells (Nurnberger and Scheel, 2001; Greenberg and Yao, 2004). For example, in pepper leaves, transient expression of calmodulin (Ca^{2+} binding protein) was observed to trigger a cell death-associated HR response following infection (Choi et al., 2009).

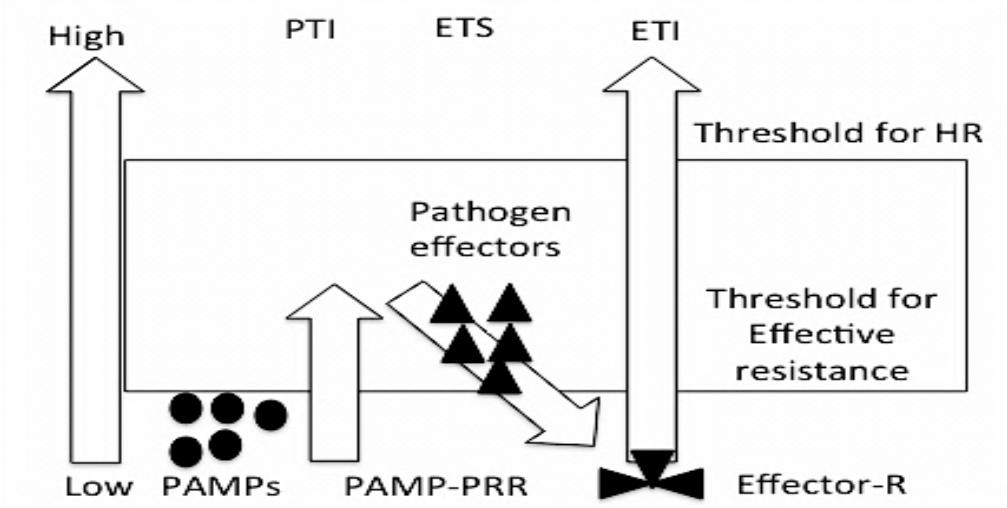


Figure 1. ETI and PTI regulation of plant defense responses

PTI is activated after recognition and perception of PAMPs (circles) PRRs. Following, the defense response passes a threshold for effective and maximum resistance. Subsequently, the surviving pathogens secret effector proteins (triangles) to suppress PTI, resulting in the activation of effector-triggered susceptibility (ETS). Lastly, the recognition of effector proteins by R proteins activates ETI (an amplified form of PTI), which directs a threshold for HR. This figure is adapted from Jones and Dangl (2006), with modifications.

2.1.3 Plant hormones associated with disease resistance

Plant hormones regulate a variety of developmental processes and responses to biotic and abiotic stresses. For plant disease resistance and immune responses, hormones such as Salicylic acid (SA), jasmonic acid (JA), and ethylene (ET) play a central role. Additionally, other plant hormones associated with developmental processes are characterized, such as abscisic acid (ABA), auxin, gibberellin (GA), cytokinin (CK), and brassinosteroid (BL). These hormones have also been recognized as regulators of plant immunity (Denance et al., 2013). Plant hormones work together in a dynamic and complex network to govern and balance the environmental and developmental signals tasked with negating or limiting defense-associated responses (Denance et al., 2013). Biotic and abiotic stress signals can alter hormonal sensitivity or concentrations, leading to a new robust response by plants (Pieterse et al., 2009).

Phytohormonal actions greatly direct and transmit early signaling events that occurr after pathogen infection into an efficient defense response. According to the type of pathogen attack incurred, different signals are sent from the infected area for synthesis of appropriate hormones at the site of infection (De Vos et al., 2006). Subsequently, a systemic defense response is triggered in distal tissues to protect the undamaged parts of the plant. This long-lasting response is referred to as systemic acquired resistance (SAR). Upon SAR induction, specific *PR* genes are activated, and SA is accumulated (Mishina, and Zeier, 2007; Van loon et al., 2006). Induced systemic resistance (ISR) is regulated by JA and ET pathways, and associated with the priming of defense responses against herbivores sensitive to JA (Pozo et al., 2008; Van Wees et al., 2008).

SA, JA, and ET have their own distinct functions. JA and ET are involved in resistance responses to necrotrophic pathogens, while SA is associated with biotrophic resistance (Robert-Seilaniantz et al., 2007). The interaction between necrotroph and biotroph resistances is antagonistic; induction of one weakens the other (Feys and Parker, 2000; Kunkel and Brooks, 2002). In tobacco plants, the expression of salicylate hydroxylase was shown to result in down-regulation of SA, although this did not affect JA during HR (Mur et al., 2006). Furthermore, in *Arabidopsis* plants, upon infection with a biotrophic oomycete pathogen and subsequent SA pathway activation, JA-related defense genes were shown to be suppressed, further confirming the antagonistic relationship between SA and JA. The exogenous application of SA is highly associated with down-regulation of JA-responsive genes such as *PDF1.2* (Koornneef, et al., 2008).

Although the majority of publications to date emphasize the antagonistic relationship between SA and JA, synergetic interactions have also been described. For

example, Truman et al. (2007) demonstrated that JA is associated with SAR establishment. In their work, they showed that JA biosynthesis activation and its related genes were critical for SAR establishment and resistance against *Pst*DC3000. In other work, conducted by Mur et al. (2006) using *Arabidopsis* plants, after lower concentrations of SA and JA treatments were applied, a synergistic effect was observed on the JA- and SA-mediated genes *PDF1.2* and *PR1,* respectively.

In the case of ET, its synergistic effect with SA has been well-demonstrated. For example, in tobacco ET-insensitive plants, ET has been shown to be critical for SA-mediated SAR after pathogen infection (Verberne et al., 2003). For *Arabidopsis* plants, ET has been shown to amplify the SA effect, leading to *PR1* activation (De Vos et al., 2006).

Abscisic acid (ABA) is yet another hormone important in developmental processes and disease resistance in plants (Asselbergh et al., 2008; Mauch-Mani and Mauch, 2005). ABA negatively regulates JA/ET pathways, since the expression of associated genes and JA biosynthesis is weakened by ABA (Adie et al., 2007; Flors et al., 2008). Additionally, ABA has also been shown to have an antagonistic relationship with SA; in work conducted on *Arabidopsis* plants, SAR induction resulted in suppression of ABA-associated genes (Yasuda et al., 2008). Similar to ABA, the hormone auxin is also involved in developmental processes. Auxin has been shown to interact with SA and JA/ET pathways in antagonistic and synergistic manners, respectively (Chen et al., 2007; Liu et al., 2006; Nagpal et al., 2005; Wang et al., 2007).

Gibberellin (GA), a hormone that regulates plant growth through growth-repressing DELLA proteins, is also associated with pathogen resistance in plants. Navarro et al. (2008) showed that mutations of DELLA proteins led to an enhanced

resistance to biotrophic pathogens, suggesting that GA-related DELLA protein degradation is associated with enhanced SA signaling. Nakashita et al. (2003) established that brassinosteroids are also important in the developmental processes of plants, as well as their disease resistance mechanisms. In their study, exogenous application of BR activated a broad-spectrum disease resistance signal. The authors demonstrated that BR hormone is associated with plant protection against biotrophic pathogens, independently of SA interactions.

2.1.4 Induced Systemic Resistance (ISR)

Plants have developed induced and constitutive networks of defense systems that help them survive against pathogen invasion. Induced disease resistance is a "state of enhanced defensive capacity" activated by particular environmental signals that results in activation of innate defense mechanisms to protect plants against pathogen invasion (Choudhary and Johri, 2007). Typically, JA/ET pathways govern ISR. ISR is usually triggered by plant growth-promoting rhizobacteria (PGPR), which enhance disease resistance in plants (Van Loon and Glick, 2004). Generally, rhizobacteria are beneficial to their host plant and do not cause obvious damage or necrosis; thus, the elicitor that they produce to induce systemic disease resistance is different from the elicitors produced by pathogens (Van Loon and Glick, 2004).

2.1.5 Systemic Acquired Resistance (SAR)

Systemic acquired resistance (SAR) is described as a broad-spectrum, long lasting disease resistance mechanism that is induced in plants by biotrophic pathogens. SAR induction results in activation of defense responses both at the site of infection and also in the distal tissues of plants, which are primed to trigger faster and

stronger defense responses upon pathogen challenge (Conrath, 2011). This phenomenon was first identified in 1933 by Chester, who noticed that local pathogenic infections caused immunization of distal tissues in plants. In 1966, Ross confirmed this induced response, naming it "systemic acquired resistance" (SAR). He observed that tobacco plants infected by tobacco mosaic virus (TMV) caused induced resistance in uninfected tissues. SAR is usually activated upon avirulent pathogen infection after ETI. Interestingly, Luna et al. (2012) showed that SAR immunity could also be transferred to the next generation of plants. Jenns and Kuc (1979) suggested that the long lasting responses of SAR take place in phloem. This theory was later confirmed in 1985 by Tuzun and Kuc: the authors demonstrated that the SAR long-distance signal was interrupted in tobacco plants upon removal of phloem post pathogen infection in the stem.

SAR is associated with accumulation of SA and increased expression levels of the *pathogenesis related 1* (*PR1*) gene (Denancé et al., 2013). SA accumulation is critical for SAR induction and the transmission of mobile signals. In a study involving *Arabidopsis ICS1* mutant plants lacking isochorismate synthase 1enzyme, disruption of SAR activation was observed to occur (Wildermuth et al., 2001). As an SA receptor, the co-activator NPR1 (Non-Expresser of PR gene1) is essential for SA-mediated SAR induction (Wu et al., 2012). Until the early 1990s, reports by different groups suggested that SA was actually the mobile signal during SAR; however, Vernooij et al. (1994) provided experimental evidence that confirmed SA is not the long-distance signal in SAR.

2.1.5.1 Long distance mobile metabolites in SAR

Different mobile signals have been identified as responsible for transmitting

SAR responses from the site of infection to the remaining plant tissues. One of them is methyl-salicylic acid (MeSA), a volatile derivative of SA (Dempsey and Klessig, 2012). Initially, Park et al. (2007) showed that in tobacco plants infected with TMV, MeSA levels increased both at the site of infection and in distal tissues. Furthermore, they showed that RNAi silencing of the *SAMT1* (*SA-Methyl transferase1*) gene, which codes for MeSA synthesis, resulted in failure in SAR induction. MeSA in distal tissues is hydrolyzed by esterase enzyme. This results in the production of SA. Then newly synthesized SA, together with MeSA, initializes the downstream defense-signaling activation mechanism (Dempsey and Klessig, 2012).

Suppressor of fatty acid desaturase deficiency 1(sfd1) mutants have been shown to down-regulate SAR induction upon abolishment of SA accumulation and *PR1* transcript levels following pathogen inoculation in *Arabidopsis* plants. *Sfd1* mutant plants failed to synthesize glycerol-3-phosphate (G3P) (Chaturvedi et al., 2012; Nandi et al., 2004). Furthermore, Chanda et al. (2011) showed that GP3 levels were high both at the site of infection and in distal uninfected tissues. Also, they demonstrated that application of G3P to *sfd1* mutant plants resulted in SAR re-establishment. Based on these results, G3P was suggested to be involved in SAR mobile-distance signalling. However, upon application of ^{14}C-labeled G3P to the local site of infection, a cease in restoration of G3P in distal uninfected areas occurred. Therefore, it was concluded that G3P is not the mobile SAR signal (Chanda et al., 2011).

Another metabolite that has been thought to be the SAR mobile signal is dehydroabietinal (DA). SA accumulation and *PR1* activation in systemic untreated leaves were observed upon local application of DA to *Arabidopsis*, tomato, and tobacco leaves. However, in *npr1* mutant plants, SAR induction by DA was not

observed, confirming that the DA role is upstream of the SA signalling (Chaturvedi et al., 2012)

Similar to DA, Azelaic acid (AzA), when applied locally, resulted in systematic disease resistance priming. In contrast to MeSA and DA, AzA treatment to *Arabidopsis* leaves did not induce SA accumulation and *PR1* transcription (Jung et al., 2009). Upon joint application of AzA and pathogen to leaves, SA accumulation and *PR1* expression were shown to be stronger and faster than the observed results in plants that were only inoculated with pathogen. These results suggest that AzA is involved in priming defences rather than SAR mobile signalling (Jung et al., 2009).

Genetic studies previously showed that SAR induction did not take place in *Arabidopsis dir1* (*Defective in induced resistance1*) mutant plants. However, SA accumulation, ETI, and PTI signalling responses were not affected in *dir1* mutant plants; only *PR1* expression was compromised (Chaturvedi et al., 2008; Maldonado et al., 2002). DIR1 codes for a lipid transfer protein that has been shown to contain a protein-protein interaction domain. As such, DIR1 may be an important gene involved in interactions with other proteins, such as DA and AzA, for transmission of SAR mobile signals (Shah and Zeier, 2013).

In sum, numerous metabolites have been suggested to play a role in SAR mobile signalling. However, rather than a single metabolite, it may be possible that mobile signalling is comprised of multiple proteins and lipids that allow for the transmission of a signal (Gao et al., 2015).

2.1.5.2 SA biosynthesis and SAR

Synthesis of SA first occurs in the chloroplast; subsequently, SA is transported to the cytosol by the EDS5 protein. Interestingly, in work performed by Serrano et al.

(2013), mutation in EDS5 led to SA biosynthesis abolishment within the chloroplast. SA synthesis is regulated by the SARD1 (SAR Deficient 1) and CBP60g (calmodulin binding protein 60g) proteins. SA synthesis is understood to occur upon binding of these two proteins to the ICS1 promoter (Truman and Glazebrook, 2012). Zhang et al. (2010) demonstrated that in *Arabidopsis* cbp60g sard1 double-mutant plants, activation of ICS1 and subsequent SA production was inhibited, further confirming the critical roles of these two proteins in SA biosynthesis.

SA biosynthesis occurs through a shikimic acid pathway that creates two separate sub-pathways, both resulting in SA synthesis. These sub-branches are referred to as the isochorismate synthase (ICS) and the phenylalanine ammonia-lyase (PAL) pathways, both of which use chorismate as their common precursor (Figure 2) (An and Mou, 2011; Dempsey et al., 2011; Singh et al., 2013; Yu et al., 2010).

In the PAL pathway, PAL enzymes are activated upon pathogen challenge, converting phenylalanine (Phe) to trans-cinnamic acid (Figure 2). Four PAL isoforms exist within the *Arabidopsis* genome. In different studies, treatment of plants with a PAL inhibitor resulted in a reduction of SA accumulation, increased susceptibility to pathogen infections, and failure to activate a SAR response (Huang et al., 2010; Mauch-Mani and Slusarenko, 1996; Pallas et al., 1996).

The ICS sub-branch is associated with chorismate conversion to isochorismate through ICS, and conversion of isochorismate to SA via isochorismate pyruvate lyase (IPL) (Figure 2). There are two ICS isoforms in the *Arabidopsis* genome (Garcion et al., 2008; Strawn et al., 2007). Similarly to results obtained in the abovementioned PAL pathway studies, ICS1 mutant plants failed to induce SAR responses (Wang et al., 2014; Wildermuth et al., 2001). Together, these results suggest that SA synthesis is essential for SAR induction from both PAL and ICS pathways. However, the

factors that regulate the employment of the ICS and PAL pathways for SA production have yet to be characterized (Gao et al., 2015).

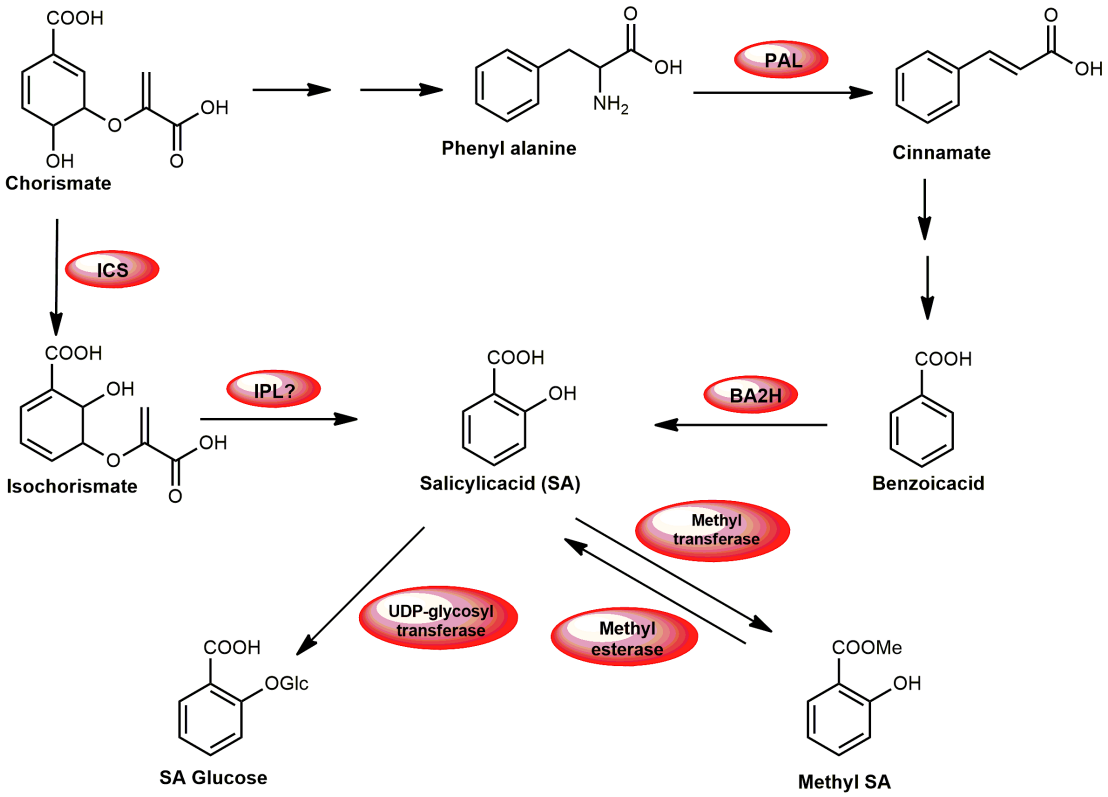


Figure 2. Salicylic acid (SA) biosynthesis in plants

ICS, isochorismate synthase; IPL, isochorismate pyruvate lyase; PAL, phenylalanine ammonia-lyase; BA2H, benzoic acid 2-hydroxylase. This figure is adapted from Gao et al. (2015) with modifications.

2.1.5.3 SA derivatives and plant defense

The majority of the synthesized SA found in plant cells is stored in the form of biological inactive derivatives. This process occurs through mechanisms such as methylation and glycosylation, among others. Examples of such derivatives are methylsalicylate (MeSA), and SA glucose ester (SGE) (Vlot et al., 2009; Dempsey et al., 2011). Similar to results obtained for SA, when tobacco and cucumber plants were

infected with pathogen, MeSA was found to accumulate in the phloem of leaves (Métraux et al., 1990; Park et al., 2007). However, Wu et al. (2012) revealed that MeSA did not interact with NPR1 with the same affinity as SA.

Furthermore, members of the benzoic acid family, including 3-hydroxy benzoic acid (3OH-BA), 4-hydroxy benzoic acid (4OH-BA), and benzoic acid (BA), have been identified in plants as SA analogues. The plant benzoic acid family is comprised of C6-C1 aromatic carboxylic acids associated with regulation of biotic and abiotic responses within plants. Biosynthesis of plant BAs starts at the shikimate pathway (Widhalm and Dudareva, 2015). This pathway is comprised of a series of seven metabolic chain reactions that occur in the plastid, where phosphoenolpyruvate and D-erythrose 4-phosphate are converted to chorismate. Chorismate serves as the precursor for the aromatic amino acids (AAA) L-phenylalanine (Phe), L-tyrosine (Tyr), and L-tryptophan (Trp) (Maeda and Dudareva, 2012; Tzin and Galili, 2010; Wildermuth, 2006).

4-OH-BA is usually referred as biologically inactive analogue of SA (Bi et al., 1995). It had not been previously shown to be associated with biotic stress responses or the pathogenesis-related (PR) proteins synthesis till Horvath et al. (2007) demonstrated the role of 4-OH-BA in increasing abiotic stress tolerance. In their work, administration of 4-OH-BA was shown to enhance the freezing tolerance of spring wheat and increase the drought tolerance of winter wheat. 4-OH-BA is formed through the phenylpropanoid pathway (Löscher and Heide, 1994) as well as directly from chorismate by chorismate pyruvate-lyase (CPL) (Kohle et al., 2002; Meganathan, 2001). In the phenylpropanoid pathway, phenylalanine ammonia-lyase (PAL) converts phenylalanine to cinnamic acid (Bagal et al., 2012). The cinnamic acid is then converted to 4-OH-BA (Kohle et al., 2002; Smith-Becker et al., 1998).

4-OH-BA has been found in plants such as grapes, carrots (Sircar and Mitra, 2009), and oil palm (Chong et al., 2009). Antifungal and antimicrobial activities by 4-OH-BA have been reported in plants (Khadem and Marles, 2010). For example, synthesis of 4-OH-BA has been found to occur in carrot cells in response to fungal pathogens (Schnitzler et al., 1992). In other work, upon exposure of cucumber plants to pathogen infection, accumulation of both SA and 4-OH-BA in phloem were observed, in addition to brief PAL induction. However, injection of 4OH-BA into leaves did not activate disease resistance to pathogen nor accumulation of SA (Smith-Becker et al., 1998). Similarly, in carrot cells, treatment with elicitors enhanced PAL activity and 4-OH-BA synthesis (Schnitzler and Seitz, 1989; Bach et al., 1993).

Both 4-OH-BA and 3-OH-BA were not found to interact with NPR1 with the same affinity as SA in competition binding assays. These results were also consistent with their role in *PR1* expression induction in *Arabidopsis* plants, since both were shown to be poor activators of the *PR1* gene (Wu et al., 2012). Notably, 3OH-BA has been isolated in plants such as olives and grapefruit (Bendini et al., 2007), and found to forgo glucosylation.

Catechol, another SA derivative, is produced by the salicylate hydroxylase enzyme (SAH). Through this reaction, SA gets decarboxylated and then hydroxylated to produce catechol. When plants were treated with catechol, *PR1* expression was not induced (Bi et al., 1995). Furthermore, when challenged with pathogen, plants expressing SAH enzymes failed to accumulate SA or activate SAR. In other work, it was observed that upon catechol infiltration in tobacco plants, lesions formed by tobacco mosaic virus infection remained unchanged. However, when catechol was co-infiltrated with different SA concentrations, *PR1* expression was not reduced or changed (Friedrich et al., 1995). In transgenic *Arabidopsis* plants expressing the

bacterial salicylate hydroxylase enzyme SA accumulation was not observed to occur and plants became more susceptible to pathogen infection (Delaney et al., 1995; Gaffiney et al., 1993).

First identified by the Ciba-Geigy Company, 2,6-dichloro-isonicotinic acid (INA) is a compound that induces disease resistance responses in cucumber against fungal pathogens (Kunz et al., 1988; Metraux et al., 1991). INA was shown to induce SAR signalling when sprayed on cucumber, pear, rice, pepper, and tobacco plants (Kuc, 1982; Metraux et al., 1991; Uknes et al., 1992; Ward et al., 1991). However, the INA response in *Arabidopsis* plants is somewhat contradictory; while some groups have reported that the INA response did not induce SAR responses (Bektas and Eulgem, 2015; Knoth et al., 2009; Uknes et al., 1992), other have reported that INA did not interact with NPR1 in binding assay experiments, nor induced *PR1* expression (Wu et al., 2012). INA has also been shown to compete with SA for binding to NPR3 and NPR4 proteins, in addition to reducing the SA binding affinity to these proteins (Fu et al., 2012). Wu et al. (2012) suggested that INA might work through a different pathway than SA to activate SA-mediated SAR responses.

Benzo-(1,2,3)-thiadiazole-7-carbothioic S-methyl ester (BTH), identified by the Ciba-Geigy Company as well, is another SA derivative, similar to INA. BTH is a potent SAR inducer; it activates SAR signalling responses in the same manner as SA (Kunz et al., 1997; Oostendorp et al., 2001; Schurter et al., 1993). In *Arabidopsis* plants, SAR activation by BTH is NPR1-dependent and interacts with NPR1 specifically in an equilibrium dialysis-binding assay (Lawton et al., 1996; Wu et al., 2012). In rice plants, WRKY45 transcription is associated with BTH-mediated SAR responses. In different studies, WRKY45-knockdown transgenic rice plants showed

reduced resistance to fungal pathogen (Bektas and Eulgem, 2015; Shimono et al., 2007).

A novel SA derivative, isotianil, has been identified by Bayer Crop Science AG (Germany) in a joint project with the Sumitomo Chemical Company (Japan). Interestingly, isotianil was shown to not exhibit any antimicrobial effect against fungus and bacteria when tested on media in vitro, but when sprayed in rice fields, a significant impact on disease resistance was found to occur against similar microbes (Yoshida et al., 2013). Additionally, Ogawa et al. (2011) showed that administration of isotianil on rice plants resulted in activation of enzymes involved in plant disease resistance, including phenylalanine ammonia-lyase, lipoxygenase, and chitinase. Isotianil is considered an important plant activator, producing minimal side effects on soil residue while demonstrating maximum efficiency against rice blast.

2.1.5.4 SA signaling components during SAR establishment

The biosynthesis and accumulation of the phytohormone SA is the most critical factor in the establishment of SAR (Denance et al., 2013). For example, *Arabidopsis* transgenic plants expressing bacterial salicylate hydroxylase (nahG), which converts SA into catechol, failed to accumulate SA and activate SAR (Delaney et al., 1995). Likewise, *Arabidopsis sid1sid2* mutant plants failed to accumulate SA, and displayed an increased susceptibility to pathogen infections (Nawrath and Metraux, 1999). In addition to proteins that are involved with direct biosynthesis (ICS and PAL proteins) and transport of SA (EDS5 protein), there are other proteins such as EDS1 (Enhanced Disease Susceptibility 1), PAD4 (Phytoalexin Deficient 4), and NDR1 (Non-race-specific Disease Resistance 1) that are also associated with SA accumulation and SAR induction. Mutations in NDR1, PAD4, and EDS1 resulted in

slightly reduced SA accumulation levels, while mutations in ICS1 and EDS5 completely abolished SAR establishment (Feys et al., 2001; Ishihara et al., 2008; Lu et al., 2013). The SA-mediated pathway that leads to SAR establishment is dependent on a highly complex network comprised of different proteins. These include pathogenesis-related proteins (PR), Non-expressor of PR1 (NPR1), and TGA transcription factors.

2.1.5.4.1 Pathogenesis-Related (PR) Proteins

During SAR induction, SA accumulation results in the expression of *pathogenesis related* (*PR*) genes, both at the local site of infection and in distal tissues (Ward et al., 1991). As a result, *PR* genes are considered a marker for efficient SAR induction (Durrant and Dong, 2004). Altogether, 17 families of PR proteins have been identified. Among them, in *Arabidopsis,* the *PR1* gene is known as a SAR activation biomarker. Similar *PR1* genes have also been reported in other plant species. SA accumulation, and subsequently, SAR activation have been shown to induce the expression of *PR1* in 5 genes (Kinkema et al., 2000). Additionally, the *PR8* gene encodes for a lysozyme associated with defense responses against bacterial pathogens, while the *PR13* gene encodes for thionins involved in antifungal defense pathways (Van Loon et al., 2006).

Lebel et al. (1998) examined the *Arabidopsis PR1* gene activation by using deletion, linker-scanning mutagenesis, and in vivo foot printing of the PR1 promoter. They successfully identified that a minimum PR1 promoter length of 10 base pairs was necessary for *PR1* activation. Linker-scanning (LS) mutagenesis analysis demonstrated that mutation at around LS7 (mutation at positions –636 to –645) and in LS10 (mutation at positions –606 to –615) resulted in complete loss of inducibility to

INA (0.35mM) and SA (5mM). Deletion analysis confirmed that DNA sequences important for the induction of *PR1* gene expression were located between positions –698 and –621 upstream of the transcription initiation site of the *PR1* gene, since removal of sequences close to –621 completely abolished PR1 promoter induction. Moreover, additional elements located between positions –815 and –698 also appeared to contribute to full inducibility of gene expression by INA (0.35mM). In vivo foot printing results were consistent with in vitro deletion and linker-scanning mutagenesis results; inducible footprints were detected at around the same regions (LS7 and LS10).

2.1.5.4.2 Timing of SAR establishment

The signaling events during establishment of SAR occur at specific times and are organized to make sure that pathways responsible for plant resistance to pathogen challenges are activated early and remain active long enough to suppress the infection, preventing it from spreading to the rest of the tissues. The *PR1* gene is usually expressed six to eight hours post pathogen infection, which can be considered a fast response based on the timeline of induced disease resistance; however, based on studies of gene inducibility by other stimuli, this response is considered a relatively slow response (Horvath et al., 1998).

The genes associated with SAR establishment have been categorized into two groups: immediate-early and late genes (Horvath and Chua, 1996; Horvath et al., 1998). The expression of immediate-early genes require pre-existing TFs; conversely, late genes involve *de novo* synthesis of TFs (Horvath and Chua, 1996; Horvath et al., 1998; Uquillas et al., 2004). Twelve different genes, considered as immediate-early genes, have been identified to be induced in *Arabidopsis* between 30 minutes to 2

hours post SA treatment. These genes are involved with protein kinases and transcription factors (Blanco et al., 2005).

2.1.5.4.3 Non-expressor of pathogenesis related gene 1 (NPR1)

NPR1 has been established as the main protein responsible for defense responses associated with both local and systemic acquired resistance pathways (SAR). As such, NPR1 plays a significant role in plant immunity regulation. As demonstrated in different studies, *npr1* mutant plants, unable to contribute to SAR establishment or the expression of the SAR marker gene *PR1,* died when exposed to pathogen infection (Cao et al., 1994; Delaney et al., 1995).

Rochon et al. (2006) showed that NPR1 is a transcription co-activator; previously, it had been unknown whether NPR1 contained a DNA binding domain that would allow its direct recruitment to the PR1 promoter. The molecular mechanism of NPR1 that results in regulation of *PR1* expression has been well characterized. In a study involving unchallenged cells, the baseline SA concentration was shown to lead to repression of *PR1* transcription through the TGA2 transcription factor (Zhang et al., 2003; Rochon et al., 2006). However, after pathogen infection and SA accumulation, NPR1 was shown to induce *PR1* expression. This mechanism is activated through the formation of an enhanceosome between NPR1 and the TGA2 on the PR1 promoter, thereby avoiding the TGA2 repressor function (Boyle et al., 2009). While the particular role of SA in the formation of the enhanceosome is still unclear, the formation of the enhanceosome itself has been very well characterized.

Mou et al. (2003) showed that in unchallenged *Arabidopsis* cells, NPR1 is sequestered as homo oligomer complexes and held together by disulfide bridges formed by cysteine residues. In other work, Tada and colleges (2008) demonstrated

that the Cys^{156} residue is the critical amino acid in NPR1 oligomerization. Furthermore, they showed that the same Cys^{156} residue goes through a *S*-nitrosylated process by *S*-nitroglutathione (GSNO), resulting in oligomer formation. Wang et al. (2005) reported that while NPR1 monomers were found to be present in the nucleus of naïve cells, only small amounts could be traced. The presence and subsequent accumulation of SA in cells triggers thioredoxins to interact with NPR1; this results in a reduction of Cys^{156} residue and subsequent release of the NPR1 monomer, which then enters the cell nucleus for activation of *PR1* transcripts (Kinkema et al., 2000; Tada et al., 2008).

In regards to structure, NPR1 is comprised of an N-terminus BTB/POZ (Broad-complex, Tramtrack and Bric-abrac and Pox virus and Zinc finger) domain, an ankyrin repeat domain, a C-terminus transactivation domain, and also a nuclear localization sequence (Kuai et al., 2015). TGA2 interacts with NPR1 through its ankyrin repeat domain (Zhang et al., 1999). Boyle et al. (2009) showed that the BTB/POZ domain of NPR1 also interacts with TGA2, conceals its repressor domain, and converts the TGA2 from a repressor to an activator.

With the use of equilibrium dialysis analysis, Wu et al. (2012) showed that SA labeled with ^{14}C binds to NPR1 with the same level of affinity as any other hormone-receptor binding. Interestingly, the same two abovementioned cysteines (Cys 521 and 529) are also a critical component of the direct binding of SA to NPR1 in *Arabidopsis* plant (Wu et al., 2012). The direct binding of SA to the C-terminus domain of NPR1 activates a conformational change, which results in disassembly of the NPR1 oligomer. Consequently, the interaction between the C-terminus and BTB/POZ domains is interrupted, releasing the C-terminus transactivation domain from BTB/POZ auto-inhibition. This results in conversion of NPR1 to an active

transcription co-activator (Figure 3). Two cysteine residues (Cys521 and Cys529) located on the C-terminal domain of NPR1 have been shown to be essential for this process to take place (Rochon et al., 2006: Wu et al., 2012).

With the use of two new high-throughput screening methods, Manohar and co-workers (2015) were able to demonstrate that NPR1 is one of the candidates that bind SA. The first method involved crosslinking of the photo-reactive SA analog 4-AzidoSA (4AzSA) to proteins in an *Arabidopsis* leaf extract, followed by immuno-selection with anti-SA antibodies, then mass spectroscopy-based analysis. The second method ultilized photo-affinity crosslinking of 4AzSA to proteins on a protein microarray (PMA), followed by detection with anti-SA antibodies.

Recently, NPR1 has been discovered as a new receptor for principal phytohormone SA (Wu et al., 2012). According to these findings, NPR1 requires the transition metal copper for direct binding to SA, and as such, can be considered to be a metalloprotein. To date, NPR1 stands as the first example of a copper-binding protein acting as a transcription regulator. SA binding to NPR1 was shown to be sensitive to chelation by EDTA, which clarifies why NPR1 had not been identified as a SA receptor earlier (Wu et al., 2012). This is not the first time that the transition metal copper was shown to be important in hormonal signal pathways. Rodriguez et al. (1999) demonstrated that the hormone ethylene also requires copper to bind to its receptor, ETR1. More interestingly, mutations in the two cysteines (Cys521 and Cys529) located on the C-terminus of NPR1 were shown to disrupt the copper binding to NPR1, as well as interrupt SA recruitment to the NPR1 C-terminal domain (Wu et al., 2012).

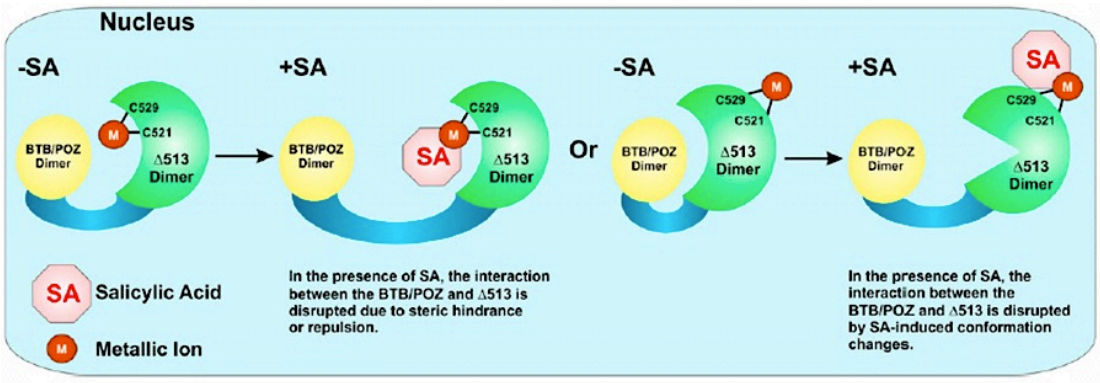


Figure 3. Model of NPR1 binding to SA and transcription co-factor activation.

NPR1 binds to SA through transition metal copper and two cysteine residues ($Cys^{521/529}$) on the C-terminus region. This results in a conformational change on NPR1- Δ513 on the C-terminus domain. The interaction between NPR1- Δ513 and the BTB/POZ domain is interrupted and the C-terminus transactivation domain is released from the auto-inhibitory domain BTB/POZ. This results in activation of NPR1 as a transcriptional co-activator regulator (Figure adapted from Wu et al. (2012)).

2.1.5.4.4 TGA transcription factors

TGA transcription factors belong to the bZIP (basic region/lucine zipper) transcription factors family, and have been found in plants, fungi, and animals (Amoutzias et al., 2008). In particular for plants, they are associated with light regulation, pathogen infection, and flower development regulation (Kirchler et al., 2010). While the leucine zipper region is important for protein dimerization, the N-terminus basic region is required for DNA binding (Jacoby et al., 2002). The N-terminal basic domain found in TGAs is abundant in proline residues, and forms a fork that enables its interaction with the related DNA sequence. Particularly, the N-terminus binds the DNA major groove when the protein is dimerized. The DNA binding domain of the bZIP family of transcription factors is conserved in Asn or Arg residues, which interact with the DNA base sequences (Miller et al., 2009).

TGA factors bind to the *as-1 (activating sequence 1)* element located in *PR* gene promoters. The *as-1* element is a 21 base-pair long DNA sequence, with two

TGACG motifs recognized by TGAs (Katagiri et al., 1989; Lebel et al., 1998). Within *Arabidopsis,* 10 TGA transcription factors have been identified, 7 of which have been shown to interact with NPR1 (Despres et al., 2000, Despres et al., 2003; Zhang et al., 1999; Zhou et al., 2000). These 7 TGAs interacting partners of NPR1 are associated with disease resistance responses based on their sequence similarity, and are clustered into 3 sub-groups. Group I includes TGA1 and TGA4, while TGA2, TGA5, and TGA6 belong to group II. Lastly, TGA3 AND TGA7 form group III (Shearer et al., 2012).

TGA1 and TGA4 have been found to either display a weak interaction with NPR1, or no interaction (Durrant and Dong, 2004), although Despres et al. (2003) clearly demonstrated that they interact with NPR1 only after treatment with SA. In their oxidizing stage, TGA1 and TGA4 tend to form disulfide bridges, which inhibit their contact with NPR1. However, when they are reduced, the formed disulfide bridges disrupt, and TGA1 and TGA4 become free to interact with NPR1 (Despres et al., 2003). Within TGA1, two cysteine residues (Cys^{260} and Cys^{266}) have been identified as important amino acids associated with the inhibition of the interaction with NPR1 before SA treatment; in work conducted by Despres et al. (2003), the interaction of TGA1 with NPR1 was observed when these two residues were mutated. In in vitro studies, Lindermayr et al. (2010) demonstrated that TGA1 is *S*-nitrosylated at the position of these Cys residues (Cys^{260} and Cys^{266}) in vitro, although they did provide in vivo experimental evidences for the existence of such reactions. This reaction leads to an increase in the affinity of the TGA1 DNA binding to the *as-1* element in the presence of NPR1.

In other work, TGA3 and TGA7 were shown to interact with NPR1 (Despres et al., 2000). They both bind the PR1 promoter in the presence of SA and NPR1

(Johnson et al., 2003). Later, Kersawani et al. (2007) showed that both TGA3 and TGA7 are mainly involved with basal resistance and are less important in expression of *PR* gene activation.

TGA2, TGA5, and TGA6 belong to class III of bZIP transcription factors in *Arabidopsis*. They are able to interact with NPR1 for activation and repression of the *PR* gene (Boyle et al., 2009; Despres et al., 2000; Rochon et al., 2006). Even though they are in the same cluster, TGA2 and TGA6 do not have similar functions. For example, Kesarwani et al. (2007) demonstrated that TGA6 is able to induce *PR* gene expression in the *tga2* mutant *Arabidopsis* plant. In the absence of SAR, both TGA5 and TGA6 bind to the *as-1* element of the PR1 promoter and interact with TGA2, which represses their function. However, after SAR induction, NPR1 interacts with them, changing them into positive regulators (Despres et al., 2003; Kesarwani et al., 2007; Rochon et al., 2006).

2.1.5.4.4.1 TGA2

The *Arabidopsis* PR1 promoter holds two cis-acting elements, referred to as linker scanning 5 (LS5) and LS7, both which contain a TGACG binding sequence. Boyle et al. (2009) showed that TGA2 binds both the LS5 and LS7 elements with similar affinity in vitro. In this work, the authors demonstrated that while a mutation in LS7 abolished *PR* gene expression, a mutation in LS5 instead induced *PR* gene transcripts in the absence of both SA and INA treatment. This confirmed that the LS7 and LS5 elements have each their own regulatory function (Lebel et al., 1998). In other work, Rochon et al. (2006) showed that TGA2 binds to the PR1 promoter both in resting and SA-induced states. However, due to the short distance sequence between LS5 and LS7 positions within PR1 promoter, researchers were unable to

make a determination as to which of these elements binds to TGA2. In this regard, Boyle et al. (2009 and 2010) hypothesized that in oligomer form, TGA2 is perhaps able to bind to both LS5 and LS7 elements, which would result in LS7 protection in the absence of SA.

In another study, when triple mutants t*ga2tga5tga6* plants were treated with INA (synthetic analog of SA) or a pathogen infection, they displayed a compromised SAR response; these results indicate that TGA2 is very important in *PR1* gene regulation. At the same time, the same triple mutants plant exhibited a high basal *PR1* transcript level, suggesting that TGA2 might act as a repressor, negatively regulating the transcription (Zhang et al., 2003). Rochon et al. (2006) demonstrated that when TGA2 interacts with NPR1 and forms an enhancesome, it acts as a transcriptional activator. In a related study, Boyle et al. (2009) confirmed that TGA2 has a dual purpose: under resting conditions, TGA2 acts as a constitutive repressor and binds to the PR1 promoter as an oligomer, repressing the expression of the *PR* gene. Conversely, in the presence of SA, TGA2 interacts with BTB/POZ and 2 ankrin repeat domains of NPR1, which results in the formation of a steady NPR1-TGA2-DNA complex. In this case, most probably, TGA2 binds to the PR1 promoter and to NPR1 as a dimer (Boyle et al., 2009).

Rochon et al. (2006) demonstrated that the only N-terminal domain of TGA2 is insufficient to repress *PR* gene activity. However, the N-terminal basic domain interacts with other domains to act as a repressor. More interestingly, Boyle et al. (2009) later determined that TGA2, through its leucine zipper domain, can form different structures such as oligomers and dimers; these conformational changes allow its binding to the LS7 element on the PR1 promoter in the absence of SA. EMSA experiments demonstrated that the full-length TGA2 protein forms two distinct

complexes with LS7 + LS5 probes; however, they observed that the TGA2 construct that lacked the N-terminus domain only formed a single complex. These results further supported previous data confirming that the N-terminal basic domain of TGA2 is critical for the formation of complexes on DNA (Boyle et al., 2009).

2.1.5.4.5 Other members of the NPR protein family in *Arabidopsis*

In *Arabidopsis,* the NPR family is composed of NPR1 plus five other NPR1-like genes, named NPR1, NPR2, NPR3, NPR4, NPR5 (BLADE-ON-PETIOLE2 (BOP2)), and NPR6 (BLADE-ON-PETIOLE1 (BOP1)) (Hepworth et al., 2005; Liu et al., 2005; Norberg et al., 2005). The BTB/POZ and ankyrin repeat domains are conserved among all NPR family proteins, while the critical cysteine residues associated with redox regulation are common in each of the NPR proteins (Hepworth et al., 2005).

2.1.5.4.5.1 NPR5 and NPR6

Among the NPR family proteins, NPR1, NPR2, NPR3, and NPR4 are very similar in structure, sharing common ancestors among them (Shi et al., 2013). However, NPR5 and NPR6 belong to a different sub-clade, sharing approximately 83.1% of their sequence identity (Table S1) (Khan et al., 2014). BOP1 and BOP2 proteins have redundant functions in plant developmental processes; in *Arabidopsis,* their functions were initially identified as floral and leaf patterning (Jun et al., 2010; Saleh et al., 2011). Recently, these two proteins have been shown to have other functions within the plant, including root nodulation structures, inflorescence, and leaf morphogenesis (khan et al., 2014). The developmental function was identified in a study involving *bop1 bop2 Arabidopsis* double mutant plants; in this study, plants

displayed leafy petioles and loss of floral abscission proteins (Hepworth et al., 2005). Through microarray studies in *Arabidopsis,* Shi et al. (2013) demonstrated that BOP2 (NPR5) was mostly expressed in pollen, hypocotyl, petiole, and shoot apex; however, weak expressions were also reported for other organs. BOP1 (NPR6) expression was found to be weak in the majority of organs, with the exception of petiole, where a high expression was found (Jun et al., 2010; Khan et al., 2014).

2.1.5.4.5.2 NPR2

NPR2 shares the highest sequence similarity with NPR1 compared with other NPR proteins. While the functions associated with NPR2 proteins have not been characterized to date (Canet et al., 2010), a poster abstract submitted to the American Society of Plant Biologists by Racki (2003) revealed interesting data concerning NPR2. In this work, NPR2 was shown to be also involved in *Arabidopsis* defense responses. The *npr2* mutant plant showed constitutive expression of *PR1* gene, in addition to unusual expression of the JA/ET pathway marker gene *PDF1.2*. In experiments involving double mutant *npr1 npr2* plants, *PR1* gene expression was completely abolished. As a conclusion, the abstract reported that NPR2 links the JA/ET signaling pathways and *PR* gene expression both in the presence and absence of NPR1. However, the information from this abstract has not yet been published in the literature.

2.1.5.4.5.3 NPR3 and NPR4

NPR1 has been well-established as a positive regulator of SAR in *Arabidopsis* plant. Additionally, two other members of the NPR protein family, NPR3 and NPR4, have also been shown to be involved in defense responses by negatively regulating

SAR (Fu et al., 2012; Liu et al., 2005; Zhang et al., 2006). Protein alignment studies have shown that NPR3 and NPR4 share approximately 34.5 and 36.0% sequence similarities with NPR1, respectively (Table S1), with the conservative region located in the BTB/POZ and ankyrin repeats (Figure S6) (Liu et al., 2005). Additionally, NPR1, NPR3, and NPR4 share four conserved cysteine residues in their BTB/POZ domain, and also five basic amino acids in their C-terminus region that might be associated with their nuclear localization (Shi et al., 2013).

Through microarray analysis, Shi et al. (2013) showed that expression of these three proteins takes place at various locations within the *Arabidopsis* plant. While NPR1 expression was only found to occur in leaf tissue, NPR3 expression was observed in young flowers, and NPR4 expression was only detected in roots and mature siliques of flower. A more detailed analysis revealed that interactions between TGA2 and the proteins NPR3 and NPR4 occurred in the nucleus (Zhang et al., 2006). However, contradictory reports have been published regarding NPR1 localization. For instance, NPR1 has been reported as being localized both in plant nucleus (Kinkema et al., 2000) and in cytoplast (Després et al., 2000).

Liu et al. (2005) showed that disease resistance responses were compromised in *npr4-1* mutant *Arabidopsis* plants after fungi and bacterial infection. On the other hand, another study, conducted by Zhang et al. (2006), observed that *npr4-3* and *npr4-2* mutant plants were not affected by bacterial pathogen, and that *npr3-1npr4-3* mutant plants showed more resistance to infection. Fu et al. (2012) further confirmed the results obtained by Zhang et al. (2006) by demonstrating that the SAR signaling responses in single mutant *npr3* or *npr4* did not differ greatly from the *Arabidopsis* wild-type plant. Additionally, they observed that the *npr3npr4* double mutant plants displayed increased resistance in both basal and SAR-induced levels.

Shi et al. (2013) also observed that when inoculated with a bacterial pathogen, the *npr3-3* mutant displayed similar bacterial growth to wild-type plants. These results were consistent with previous data; however, their findings also revealed that transgenic *Arabidopsis* plants that exhibited an overexpression of NPR3 were more susceptible to bacterial infection. The inconsistencies in findings between these groups may be attributed to the use of different mutant alleles of NPR3 and NPR4, although all findings confirmed the same role for both NPR3 and NPR4 in plant immunity (Kuai et al., 2015).

In 2012, Fu and colleagues proposed a model describing the interaction between NPR1, NPR3, and NPR4 with respect to SA and SAR signalling. Their findings demonstrated that NPR4 acts as a CUL3 E3-ligase substrate in resting cells and interacts with NPR1, resulting in NPR1 degradation targeting by ubiquitin proteasome. However, when SA is accumulated in cells and SAR is activated, the NPR4 binds SA, interrupting the interaction between NPR4/NPR1; consequently, NPR1 degradation is abolished. Conversely, NPR3 binds to SA and interacts with NPR1, which is followed by a restart in NPR1 ubiquitylation and degradation. Based on the proposed model by Fu et al. (2012), NPR4 regulates NPR1 degradation in unchallenged cells, and similar to NPR4, NPR3 is also involved with NPR1 degradation, but in SAR-induced cells.

Despite the remarkable insights provided by the model proposed by Fu et al. (2012), Kuai et al. (2015) provided clear explanations that reject the model. Foremost, the model has been critiqued owing to missing clarifications regarding the mechanisms that allow NPR3 and NPR4 to bind to SA, and as to whether any conformational changes occur, similar to any hormone receptor interaction. Secondly, contradicting the results obtained by Fu et al. (2012), where the authors offered

experimental evidence that NPR1 does not bind SA, Kuai et al. (2015) refuted these findings by presenting results from two independent studies that clearly confirmed that NPR1 interacts with SA (Manohar et al., 2015; Wu et al., 2012). Thirdly, knowing that SA binds NPR1 with high affinity through the transition metal copper, the original model failed to demonstrate that the interaction between NPR1 with NPR3/NPR4 is due to SA binding to NPR1 or NPR3/NPR4. In opposition to this model, Kuai and colleagues suggested that perhaps NPR1 would operate between NPR3 and NPR4 to direct the occurrence of any interactions among these three proteins (Kuai et al., 2015). To support this theory, the authors highlight that NPR1 is the only protein among all the three proteins that has been shown to go under conformational change similar to any other hormone receptor interaction (Wu et al., 2012). Further, *PR1* gene expression in unchallenged cells was revealed to be higher in *npr3* mutants and at about the same level in *npr4* mutants when compared to wild-type plants. However, expression levels were found to be relatively higher in double mutant *npr3npr4* plants compared to those found for single mutants of *npr3* and *npr4,* as well as in wild-type plants. These results are not in agreement with the model proposed by Fu et al. (2012); based on the model, a similar level of *PR1* expression would be expected in *npr4* single mutants and *npr3npr4* double mutants, while *PR1* induction levels should be similar in *npr3* single mutants and wild-type plants (Kuai et al., 2015).

2.1.5.4.6 NPR1 in other plant species

Homologues of *Arabidopsis* NPR1 (*At*NPR1) (Cao et al., 1998) have been identified in crop plants such as cotton (Zhang et al., 2008), rice (Chern et al., 2005; Yuan et al., 2007), soybean (Sandhu et al., 2009), banana (Endah et al., 2008; Zhao et

al., 2009), rosaceous tree (Pilotti et al., 2008), apple (Malnoy et al., 2007), and grapevine (Henanff et al., 2009). Additionally, overexpression of *At*NPR1 in rice, tomato, wheat, and apple was shown to increase disease resistance responses to fungal and bacterial pathogens (Chern et al., 2001; Lin et al., 2004; Makandar et al., 2006; Malnoy et al., 2007).

The NPR1 gene has been cloned and characterized in Sugarcane (*Saccharum spp.* hybrids) (*Sc*NPR1). It has been shown to share an 83% sequence identity with NPR1 from maize (*Zm*NPR1), as well as display a very close relationship with NPR1 from banana (*Md*NPR1). As in any other NPR1 protein, it contains a BTB/POZ zinc finger domain, ankyrin repeats, and basic amino acids in its C-terminal region for nuclear localization. In a study conducted by Chen et al. (2012), the expression level of *Sc*NPR1 was induced after SA treatment and fungal pathogen challenge; these findings suggest that *Sc*NPR1 might also be associated with SA-mediated defense signaling pathways.

Furthermore, the NPR1 in rice (*Oriza sativa)*(OsNPR1) plants has also been shown to be involved in defense signaling responses. Overexpression of *At*NPR1 in transgenic rice plants was shown to enhance defense responses against the rice blight bacterial pathogen. Similar to *Arabidopsis,* TGA transcription factors of rice were shown to interact with NPR1. Although high levels of endogenous SA have been observed in rice plants under a resting state (Silverman *et al.,* 1995), bur rice and *Arabidopsis* NPR1 proteins were revealed to conserve responses to disease resistance (Chern et al., 2001; Sugano et al., 2010).

Two copies of NPR1 have been identified in soybean (*Glycine max*); *Gm*NPR1 and *Gm*NPR2. Overexpression of these proteins in transgenic soybean plants showed enhanced disease resistance against pathogen infection and *PR1*

expression induction. Similar to *At*NPR1, expression levels of *Gm*NPR1 and *Gm*NPR2 were induced after SA treatment, suggesting that SA-mediated SAR signaling is also conserved in soybean plants (Sandhu et al., 2009).

NPR1 from cotton (*Gossypium hirsutum*) (*Gh*NPR1) displayed 52.98% and 52.32% sequence similarity with *Arabidopsis* and tobacco NPR1 proteins, respectively. Similar to the findings regarding *Os*NPR1a, the same BTB/POZ domain, ankyrin repeats, and nuclear localization signals were found in *Gh*NPR1. The mRNA levels of *Gh*NPR1 were found to be induced in response to SA administration and pathogen challenge (Zhang et al., 2008).

Two copies of NPR1-like genes have been found in banana (*Musa*), *M*NPR1A and *M*NPR1B, with both genes shown to be very sensitive to SA treatment and pathogen infection. However, the *M*NPR1B responses were found to be much stronger compared with the responses obtained for the *M*NPR1A gene. These results suggested that the SA-mediated SAR pathways might be also conserved in banana, as observed in other plant species (Endah et al., 2008).

Although the cysteine residues ($Cys^{521}/^{529}$) found to be critical for SA perception by NPR1 in *Arabidopsis* are not conserved among plant species such as rice, tobacco, cacao, and cotton (Wu et al., 2012), the NPR1 proteins in these crops have been shown to be sensitive to SA treatments. These findings suggest that metal binding and SA perception in these crops occurs through a process that involves other amino acids with electronegative property; as such, the NPR1-SA binding activity would be expected to occur in a similar way among plant species, although the particulars of the connection would be different in each crop, based on the particular requirements within that crop (Wu et al., 2012).

2.2 Protein-DNA interaction

Protein and DNA are the most significant biomolecules in any living organisms; the DNA stores genetic information, while proteins are responsible for the regulation of life mechanisms. Accordingly, the interactions that occur between DNA and protein are fundamental in the regulation of many different biological processes (Rohs et al., 2010). Interactions between proteins and nucleic acids are categorized into two groups: specific and non-specific interactions. In non-specific interactions, the sequence order of nucleotides is unimportant, as long as a binding interaction occurs. Conversely, specific interactions are regulated by sequence recognition (base readout), as well as the geometries of both protein and DNA (shape readout) (Rohs et al., 2010).

Interactions between protein and DNA occur through physical forces such as electrostatic interactions (e.g., salt bridges), dipolar interactions (e.g., hydrogen bonding), and entropic effects (e.g., hydrophobic interactions, dispersion forces involving base stacking) (Rohs et al., 2010). These forces play an important role in binding of proteins to a specific or non-specific sequence of nucleic acids. The affinity and specificity of a specific protein-nucleic acid interaction can be further enhanced by protein oligomerization or the formation of multiple protein complexes, such as the transcription initiation complex (Strauch, 2001).

Protein-DNA interactions are generally regulated by an α-helix motif in the protein that inserts itself into the major groove of the DNA. This allows for recognition to take place, permitting interactions to occur with a specific base sequence through the formation of salt bridges and hydrogen bonds. Each individual protein that binds to DNA or RNA has a distinct conserved domain in its tertiary structure that contributes to the binding process. Depending on its structure, a protein

can have multiple repeats of the same binding domain. Common DNA binding domains within proteins are helix-turn-helix, helix-loop-helix, zinc fingers, and leucine zipper (Siggers and Gordon, 2013).

2.2.1 Principals of protein-DNA recognition

Interactions between protein-DNA are highly dependent on the three-dimensional structures of both interacting partners. Similar to other macromolecular interactions, the forces that contribute to the protein-DNA complex formation are quite weak; as such, the sum of multiple forces is required to achieve overall affinity (Rohs et al., 2010). The most recognizable among these forces is electrostatic force; this is due to the polyanion nature of nucleic acids and binding proteins with positively charged amino acids, such as lysine and arginine. Polar interactions are also common, including water-mediated hydrogen bonds, which play a key role in the sequence specificity of binding. In addition, the primary status of individual proteins and DNA is a very important factor in these interactions, as complex formations can cause changes in flexibility and conformation in either or both (Garvie and Wolberger, 2001).

The simplest and most direct way for proteins to recognize a specific sequence in DNA is to bind in the major groove and examine the features of the bases that are presented; this would allow the protein to make favorable contacts with the correct sequence. The minor groove of DNA is less interesting in this respect, since all four bases show hydrogen bond acceptors in similar locations (Rohs et al., 2009), hindering the protein molecules from distinguishing between bases. In 1959, Zubay and Doty proposed that the alpha (α) helix found in proteins fits very well into the major groove of B-form DNA. In contrast to the major groove of B-form DNA, the

major groove of A-form DNA is characterized by a deeper and narrower structure, and thus less accessible to binding proteins. In indirect sequence recognition (readout) by proteins, the DNA sequences are recognized based on sequence-dependent differences in flexibility and structural factors, such as groove width and particular twists between base pairs (Rohs et al., 2010).

The recognition mechanisms between protein and DNA are categorized into two major classes, including shape and base readouts, and are further subdivided into groups.

2.2.1.1 Base readout

Base readout, which refers to the process where proteins come into contact with the bases in either the major or minor groups of DNA, is one of the ways for DNA-binding proteins to gain specificity for interaction. This recognition mechanism is usually regulated by the formation of hydrogen bonds between amino acids and bases, and in some cases, through water-mediated hydrogen bonds or hydrophobic contacts. Hydrogen bonds with bases contribute to greater specificity in the major groove of DNA as compared to the minor groove. This is due to the presence of specific patterns of hydrogen-bond donors and acceptors in base pairs of DNA major groove. Proteins, on the other hand, use the helix-turn-helix domain, forming hydrogen bonds with bases in the DNA major groove. The degree of specificity achieved for protein-DNA interactions through hydrogen bonds is greatly dependent on the number of contacts between DNA bases and protein residues, and more importantly, on the uniqueness of the hydrogen-bonding geometry (Rohs et al., 2010).

Highly ordered water molecules also regulate the precise readout of bases in the protein-DNA recognition process. For example, it has been reported that in

interactions occurring within the Lac repressor, the protein-DNA complex keeps a large amount of its hydration when it binds non-specifically; however, the same occurrence has not been observed for specific interactions. (Kalodimos et al., 2004). Thus, it can be assumed that water-mediated hydrogen bonds occurring in the major groove of DNA contribute to the specificity of the base readout, since they indicate the positions of hydrogen-bond donors and acceptors. However, this cannot be applied for water-mediated hydrogen bonds in the minor groove, as the exact formations of the hydrogen-bond donors and acceptors have yet to be identified (Jayaram and Jain, 2004).

2.2.1.2 Shape readout

For the majority of DNA-binding proteins, hydrogen-bond base readouts and hydrophobic contacts are not enough factors to specifically recognize their interacting DNA partners. Other factors, such as sequence-dependent DNA structure and deformability, also contribute to recognition specificity mechanisms (Rohs et al., 2009).

2.2.1.2.1 Minor and Major groove shapes

Narrow minor grooves present great mechanisms for recognition of DNA shape sequence-specificity due to their enrichment in the AT sequences and increased negative electrostatic potential, which is in turn recognized by arginine residues (Rohs et al., 2009). Similar to minor grooves, the sequence-dependent major groove of DNA is also applied as a readout mechanism among DNA-binding proteins. Naturally, in DNA, major and minor groove geometries are associated with each other. For example, in humans, the regulatory factor hRFX1 is a winged helix-turn-helix protein

that recognizes the major groove of DNA via its β-hairpin wing. In fact, the hRFX1 protein localizes its H3 helix over the minor groove from the point where a single lysine residue contacts the groove (Gajiwala et al., 2000). Consequently, the minor groove widens, ultimately resulting in a narrowing of the major groove and in an enhancement in the shape of the major groove.

2.2.1.2.2 DNA kinks

DNA kinks occur when the linearity of the helix is sharply broken, usually due to the un-stacking of a flexible base pair step. Kinks provide a great binding specificity opportunity by enhancing structural confirmations, which result in protein-DNA contact improvement (Sierk et al., 2001). An example of kink association with protein-DNA contact has been reported for the DNA-binding site of the catabolite activator protein (CAP) (Figure 4A) (Lawso et al., 2004). Drastic kinks resulted in DNA bending to approximately 90◦ around the CAP protein (Schultz et al., 1991). Generally, due to weaker stacking interactions, kinks are stabilized by insertion of protein side chains. For example, in the case of the Lac repressor, its specific DNA-binding site adopts to the protein through the formation of a kink of 36◦ at its central CPG step (Figure 4B) (Kalodimos et al., 2004).

2.2.1.2.3 Bent DNA

DNA bending has a significant role in nucleic acid sequence recognition. In addition to van der Waals and hydrogen bonding interactions, the sequence bendability of DNA also contributes to sequence discrimination by providing the specific conformation needed for protein binding at a lower free-energy cost. DNA bending has been observed in the majority of DNA-protein interactions, with the

degree of bending ranging from minor kinks to a 360-degree turn in certain cases (Figure 4C) (Steitz, 1990). Such bends are usually associated with changes in the groove width and depth as well as alterations in the sugar-phosphate backbone. DNA bending can be categorized into two groups: the first group includes bending processes that occur locally at a single site, while the second group is comprised of local changes at the sequential base steps that bend the helix (Xiong, and Sundaralingam, 2001).

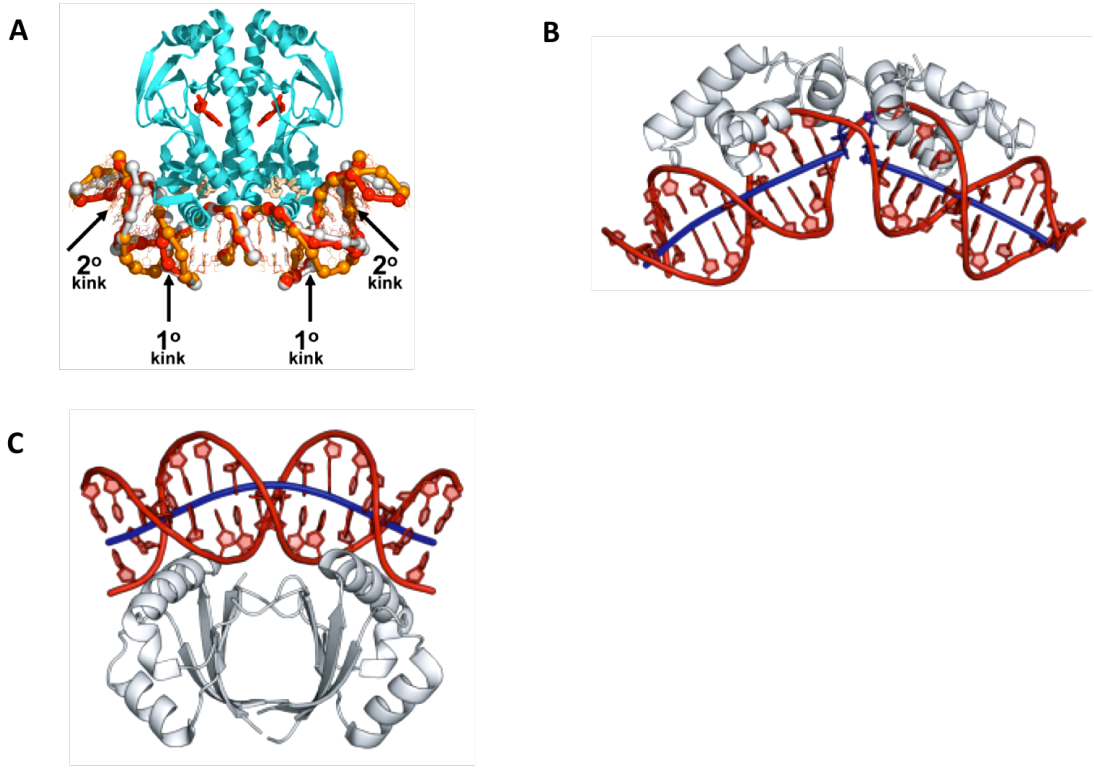


Figure 4. Readout mechanisms based on kinking and bending in protein-DNA complexes

A. Structure of DNA binding by CAP. The interaction occurs through a substitution in the primary kink site. The kink widens minor groove and facilitates the interaction.
B. Structure of DNA binding by *Lac repressor*. The calculated helix axes for both sides of the kink (blue lines) show an abrupt change in the helix trajectory caused by the kink.
C. HPV-18 E2 protein binds to DNA through bending over a stretch of the helix. The smooth bend is pictured by the helix (blue lines).
(Figures are adapted from Lawson et al. (2004) and Rohs et al. (2010) with modifications).

2.2.2 DNA-binding proteins

Sequence-specific DNA binding proteins are associated with different DNA regulatory events, such as replication, recombination, repair, and transcription. Transcriptional regulators comprise the largest class of DNA binding proteins, which bind DNA to start mRNA transcription. They also regulate different cell processes including growth, development, and differentiation. Transcription factors bind DNA by binding to a specific DNA sequences located upstream of the gene promoter, which either activates or represses related genes by interaction with RNA polymerase II or other co-factors bound to the promoter (Luscombe et al., 2000).

2.2.2.1 Helix-turn-helix (HTH) proteins

The Helix-turn-helix (HTH) is a DNA recognition motif found in bacteriophage, prokaryotes, and eukaryotes. The HTH transcription factors from bacteriophage and bacteria are normally homodimers, while the eukaryotic ones are monomeric. The basic HTH fold is composed of three core helices that form a right-handed helical bundle with a tight turn between helices 2 and 3 (Luscombe et al., 2000). All HTH motifs include a hydrophobic core at the interface of the three helices that help stabilize the motif. Helix 3, also referred to as the DNA recognition helix, inserts itself into the major groove of the target DNA with its side chains contacting the nucleotide bases, forming both specific base and sugar-phosphate backbone contacts (Figure 5A) (Brennan and Matthews, 1989; Rohs et al., 2010).

The *Cro* and *Repressor* family of proteins (Figure 5D) comprises one of the HTH protein groups. These proteins are involved in lysogenic/lytic growth-switch mechanisms in bacteriophages that contribute to transcriptional regulation at associated operons. *Cro* and *Repressor* proteins bind to operons with various degrees

of affinity. Each operon is comprised of approximately 14 base pairs in length, with four conserved bases at the end. The recognition helix from the HTH motif links with the base ends in the major groove of DNA (Ohlendorf et al., 1982).

2.2.2.1.1 Winged HTH proteins

Winged HTH proteins, described as having a third α helix and a β sheet, are considered an extension of the HTH family and believed to be factors for the DNA-binding motifs. Similar to the structure found in HTH motifs, the recognition helix of winged HTH proteins is responsible for the binding, while secondary structure features offer extra contacts with the DNA backbone (Luscombe et al., 2000; Rohs et al., 2010).

2.2.2.2 Zinc-coordinating proteins

Zinc finger proteins are very well-characterized proteins involved in different plant activities such as growth, development, biotic, and abiotic stress resistance (Furtado et al., 2011; Giri et al., 2011). Any functional, small folded domain that coordinates one or more zinc ions to stabilize the protein structure is called a zinc finger protein (Laity, 2001). Zinc finger domains were originally producers of regulator proteins (Klug and Rhodes, 1987). In R proteins, zinc finger domains are located both in the C and N-terminals (Eulgem and Somssich 2007; Gupta et al., 2012).

There are more than ten different subfamilies of zinc-binding domains. They are classified based on their type, cysteine and or histidine, as well as the arrangement of amino acids that regulates tetrahedral zinc coordination, which influences the overall fold of the domain. All classes of zinc-binding domains share a common

feature; during DNA-binding, zinc ions are used as a structural element to maintain the folding of a rather short polypeptide chain, which would not otherwise fold correctly due an insufficient number of hydrophobic residues necessary to stabilize the core of the domain (Figure 5B) (Laity et al., 2000; Luisi, 1992).

Among the zinc-binding domains that recognize the core DNA segments are the classical Cys_2His_2 zinc fingers and Cys_2Cys_2 GATA-1 like domains. The classical Cys_2His_2 zinc fingers are small motifs considered to be the largest subgroup of zinc-containing DNA-binding proteins. They are found both in prokaryotes and eukaryotes. The typical sequence motif for a single classical zinc finger is X2-Cys-X2-4-Cys-X12-His-X2-8-His, which coordinates a single zinc ion. Approximately 30 residue repeats conglomerate to form a two-stranded anti-parallel β-sheet followed by an α-helix, which are then held together by a small conserved hydrophobic core and zinc ion. For DNA binding, the β-sheet contacts the backbone while the α-helix inserts into the major groove, making a sequence-specific side chain to base contacts (Pavletich and Pabo, 1991; Wolfe et al., 1999).

2.2.2.2.1 ββα zinc-finger proteins family

The ββα zinc-finger protein structure is composed of a short two-stranded anti-parallel β sheet with an α helix. Two pairs of conserved histidine and cysteine residues in the α helix and a second β strand organize a single zinc ion. Generally, the proteins from this group contain multiple fingers, which are wrapped around the DNA in a twisting way, binding the 2 base pairs via insertion of the α helix in the major groove of the DNA (Luscombe et al., 2000; Rohs et al., 2010).

2.2.2.2.2 Loop-sheet-helix family

The DNA-binding domain of proteins from this family include a primary loop out of the main body of the protein, followed by a small β sheet, and an α helix followed by another loop. A histidine and three cysteines in the two loops accommodate the zinc ion. The α helix of the protein binds to the major groove of the DNA. Protein p53, a transcriptional activator associated with tumor suppression, is an example of proteins included in this family (Siggers and Gordon, 2013).

2.2.2.3 Zipper type proteins

Zipper-type proteins are categorized into two families, including leucine zipper and helix-loop-helix proteins. All protein DNA bindings from this group are considered to be homodimers, with the exception of a few that function as heterodimers (Alber, 1992).

2.2.2.3.1 Basic leucine-zipper

The basic leucine-zipper (bZIP) DNA-binding motif is a dimer of extended α-helices. It contains a C-terminal dimerization region stabilized by a heptad repeat of hydrophobic residues along one face of the helices. The helices enables a zipper-like coiled-coil. The N-terminal end of the helices contains more basic and polar DNA-binding parts that insert into opposite sides of the same DNA major groove. In the basic part, one side of the helices contacts the DNA base functional group, while the neighboring residues contact the phophodiester backbone (Figure 5C). These bZIP proteins are only found in eukaryotes (Ellenberger et al., 1992; Gordan et al., 2013).

2.2.2.3.2 Basic Helix-loop-helix

The basic helix-loop-helix (bHLH) domain has similar characteristics to the bZIP domains. For example, the bHLH domain also contains a C-terminal coiled-coil dimerization region and an N-terminal basic region, which contacts the DNA major groove. The difference between the bHLH and bZIP domains stems from a different configuration in the N- and C-terminal segments of the bHLH domains. These segments are not part of the same helix, but two separate helices connected by a loop and helical segments of the dimer, which originate from a four-helix bundle at the interface between the dimerization and the DNA recognition segments (Gordan et al., 2013; Siggers and Gordon, 2013).

2.2.2.4 β sheet proteins

Proteins from this group use β-strand structures for DNA recognition and binding. The TATA box-binding protein family is an example from this group (Figure 5E) that uses a wide β sheet for DNA binding. TATA box-binding proteins are very important factors for transcription initiation interactions. The structure of the TATA box-binding protein is comprised of two identical domains, with a ten-stranded anti-parallel β sheet linking the domains and covering the DNA minor groove (Burley, 1996 and Tateno et al., 1997).

2.2.2.4.1 β-hairpin/ribbon proteins

Proteins from this group use two or three-stranded β sheets or a hairpin motif to bind to DNA, and contain β sheets that are smaller than the ones that TATA box-binding proteins use. Different families of proteins have been characterized from this group, including the MetJ repressor (Figure 5F) (Raumann et al., 1994), Arc repressor, and T-domain families (Somers and Philips, 1992). Each individual protein

structure is composed of a helical bundle and a single β strand. The strands form an anti-parallel sheet and bind to the major groove of DNA.

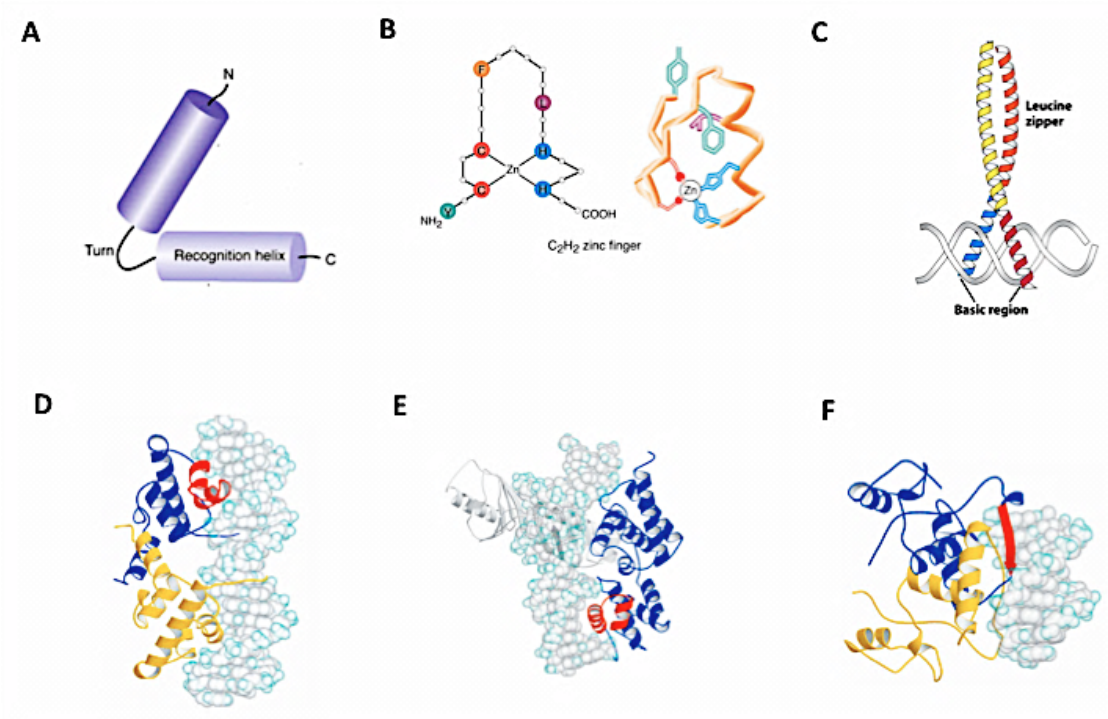


Figure 5. Schematic representing examples of DNA binding motifs and their DNA binding proteins

A. Helix-turn-helix (HTH) motif. Nonpolar contacts between these helices stabilize the motif. Residues on the recognition helix form hydrogen bonds with the bases in the major groove.
B. Zinc finger motif. A zinc atom is conjugated with 2 cysteines (C) and 2 histadines (H) of a small part of a polypeptide chain, forming a finger-like shape.
C. Basic leucine zipper motif. The leucine zipper is a stretch of amino acids, mainly comprised of leucine residues. It provides a dimerization motif, which is stabilized by hydrophobic residues.
D. The *Cro* and *Repressor* family. The DNA-binding motif is shown in red. The protein binds as a dimer; one monomer is colored blue and the other yellow.
E. The TATA box-binding family. The DNA binding motif is shown in red. One monomer is colored blue.
F. Met repressor β hairpin/ribbon proteins. The DNA binding motif is shown in red. One monomer is colored blue.
(These images are adapted from Luscombe et al. (2000) and Alberts et al. (2002), with modifications).

2.2.3 Complexities in protein-DNA recognition

The DNA-binding specificity is encoded at multiple levels during DNA-protein interactions. At each level, difficulties can arise due to variations in the readout mechanisms.

2.2.3.1 Transcription factor (TF)-specific preference

The binding sites of transcription factors regulate transcription activities within the genome (Priest et al. 2009). DNA sequences, usually 6 to 10 base pairs in length, are located in the flanking regions of genes, and are usually the sites for binding of transcription factors (Wray et al. 2003). Promoter regions of genes have multiple sites for potential transcription factors binding, which can result in expression or repression of specific genes (McKenna and O'Malley 2002). Closely associated transcription factors often display both common and TF-specific binding preferences (Berger et al., 2008).

Results from a study of mouse transcription factors have demonstrated that among 105 TFs, even proteins with high amino acid sequence similarity (as high as 67%) demonstrated different DNA-binding profiles. In the majority of cases, individual members of a DNA binding domain protein family were found to possess identical highest affinity sites; however, individual proteins within the same family had distinct preferences for lower affinity sites. For example, mouse TFs interferon regulatory factor 4 and 5 (Irf4 and Irf5) had similar preferences, with strong binding affinity for sequences containing the 5'-CGAAAC-3' site, but preferred sequences having the 5'-TGAAAG-3' for lower binding affinity. Such TF-specific preferences resulted in difficulties for DNA-binding models, since models have to precisely distinguish common binding sites, while being able to predict specific sites for each TF as well (Badis et al., 2009; Noyes et al., 2008).

2.2.3.2 Flanking DNA

DNA sequence flanking in the core can affect the binding affinity between DNA and protein and give rise to complications, even for TFs with precise core DNA sequence binding sites. For example, the transcription factor general control non-repressible 4 (Gcn4) binds to core DNA sequences that contain 5'-TGACTCA-3'. Binding measurement analysis of Gcn4 to thousands of DNA sequences having this exact sequence showed various forms of binding affinity profiles, ranging from inconsiderable binding to high-affinity binding. In work conducted to investigate the DNA-binding preferences of the yeast TFs C-promoter binding factor 1 (Cbf1) and Tye7 to genomic sequences having the E-box motif 5'-CACGTG-3', the authors demonstrated that the sequence beyond the canonical E-box motif provides binding specificity by affecting the three-dimensional structure of the DNA-binding site (Nutiu et al., 2011). Essentially, further specificity originating from sequences flanking the core motif cannot be solely explained by the extension of the core; as such, it is possible that the impact of flanking DNA is perhaps to be employed by DNA shape and not by specific DNA contacts (Gordan et al., 2013).

2.2.3.3 Multiple modes of DNA binding

Complexities arise in DNA binding for proteins that bind DNA using two or more distinct modes, including variable spacing, multiple DNA-binding domains, and alternate structural conformations. Such different modes of interactions essentially result in different DNA-binding preferences, making the implementation of simple binding models a challenging undertaking (Berger et al., 2006).

2.2.3.3.1 Variable spacing

Some proteins are able to bind to DNA motifs that consist of two half-sites. The half-sites are separated by a different number of bases. This type of binding was first demonstrated for basic leucine-zipper (bZIP) proteins that can bind overlapping or adjacent TGAC half-sites (Siggers and Gordan, 2014). Later studies have categorized the bZIP family into two subfamilies according to their protein sequences. Proteins in the first group are called Activator protein 1 (AP-1) proteins; they display a binding preference for overlapping half-sites but bind to adjacent half sites with almost equal affinity (Kim and Struhl, 1995). Proteins in the second group are called Activating transcription factor (ATF)/cAMP response element binding (CREB) proteins; these proteins exhibit a binding preference for adjacent half sites, but also bind overlapping half sites weakly (Vinson et al., 2002).

Variable half-site spacing has also been reported for nuclear receptors (NRs). For example, the peroxisome proliferator activated receptors (PPARs) can bind to response elements consisting of direct repeats of the 5'-AGGTCA-3' half site, which is separated by different length spacers (Khorasanizadeh and Rastinejad, 2001).

2.2.3.3.2 Multiple DNA binding domains (DBDs)

DNA binding proteins can have multiple, independent DBDs. This allows the protein to recognize different DNA segments by using DBDs. For example, the zinc fingers (ZF) protein ecotropic viral integration site 1 (Evil) has 10 ZF domains divided into two autonomously functioning DBDs, including an N-terminal seven ZF DBD and a C-terminal three ZF DBD, which are identified as distinct DNA motifs (Delwel et al., 1993).

2.2.3.3.3 Alternate structural conformations

Proteins with multiple DBDs recognize separate DNA motifs and bind to multiple distinct DNA regions (Ben-Gal et al., 2005). The ortholog of human TF sterol regulatory element-binding protein 1 (SREBP) is an example of such a protein. SREBP proteins are part of the basic helix-loop-helix (bHLH) family of TFs (Kim et al., 1995), which bind asymmetric sterol regulatory element 5'-ATCACnCCAC-3'. A co-crystal structure study of the DNA-binding domain of human SREBP-1a revealed that the protein binds to an asymmetric DNA–protein interface; one monomer binds the E-box half-site (5'-ATCAC-3') through protein–DNA contacts, while the second monomer recognizes the non-E-box half-site (5'-GTGGG-3') using completely different protein–DNA contacts. As such, for SREBP, residues in DBD facilitate multiple modes of DNA binding. However, proteins that are able to recognize different DNA-binding modes are involved with difficulties such as the presence of separate modes that normally associated with a necessity to apply multiple DNA binding models to show the specificity of the protein (Kim et al., 1995).

2.2.3.3.4 Cooperative binding

Cooperative binding of TFs to DNA stabilizes DNA proteins, while improving their roles for transcription of genes. This usually takes place by protein-protein interactions between bound TFs. Cooperative bindings also influence the DNA-binding distinction and result in recognition of novel binding sites. One of the ways that cooperative binding changes binding specificity is via extension of the TF binding to lower affinity sites (Wolberger, 1999). A well-studied example of such cooperative binding is the binding process that occurs between yeast TFs MATα2, MATα1, and Mcm1 (mini-chromosomal maintenance binding protein). MATα2 TF binds to DNA cooperatively with help of co-factors MATα1 or Mcm1 in order to

suppress various sets of genes. MATα2: MATα1, and MATα2: Mcm1 heterodimers bind to sites found in haploid-specific and mating-type specific promoter regions of genes through recruitment of co-repressor proteins (Carey, 1998; Siggers and Gordon, 2013).

2.2.4 Methods for studying DNA-protein interaction

DNA-protein interaction is a very important process for the majority of cellular function. To explain such complex interactions, different in vivo and in vitro techniques have been developed over the years. All these techniques have their own pros and cons and are associated with a particular purpose (Dey et al., 2012).

2.2.4.1 Chromatin Immunoprecipitation (ChIP) related techniques

Chromatin Immunoprecipitation (ChIP) is an immunoprecipitation technique used to study the specific interaction between DNA and proteins, using a specific antibody against the protein of interest. To date, only one well-established in vivo ChIP method has been developed (Solomon et al., 1988). The workflow of the ChIP experiment is illustrated in Figure 6. The first step in the ChIP experiment involves in vivo crosslinking of DNA-binding proteins to the genomic targets. Next, the protein-DNA complex is isolated from the cells, and bound DNA is sheared into shorter fragments by sonication or enzymatic digestion. Subsequently, the DNA fragments that are cross-linked with the protein of interest are immunoprecipitated (IP), using a specific antibody towards the protein of interest that specifically binds to it. Finally, the IP-enriched DNA fragments are analyzed using various techniques, such as polymerase chain reaction (PCR) and quantitative PCR (qPCR), both of which allow for a determination of whether the pre-defined DNA sequences are enriched in the

identified DNA fragments.

Other techniques, such as tiling array (ChIP-chip) and high-throughput sequencing (ChIP-seq), can also map the identified DNA fragment within the locations in the reference genome. ChIP-seq is a widespread technique that allows for precise determination of motifs at a nucleotide-level resolution (Bortz and Wamhoff, 2011; Mokry et al., 2010; Park, 2009; Raha et al., 2010; Robertson et al., 2007). Prior to sequencing, target DNA fragments are amplified to form a library. Then, the immunoprecipitated DNA of 150 to 300 base pair sizes is selected for sequencing. In the ChIP-chip technique, instead of sequencing, the immunoprecipitated DNA is bound on a microarray (Pillai and Chellappan, 2009). However, due to its low resolution, low coverage, high cost, and smaller artifact yield, ChIP-chip is not used as widely as ChIP-seq. Additionally, ChIP-seq is more often selected as the method of choice due to its ability to map histon modifications and DNA methylation (Barski et al., 2007; Li et al., 2011). Although ChIP-seq has been very well-established since its introduction, there are some drawbacks associated with this technique; the technique is burdened with antibody sensitivity and specificity, in addition to a complicated bioinformatics analysis protocol (Park, 2009).

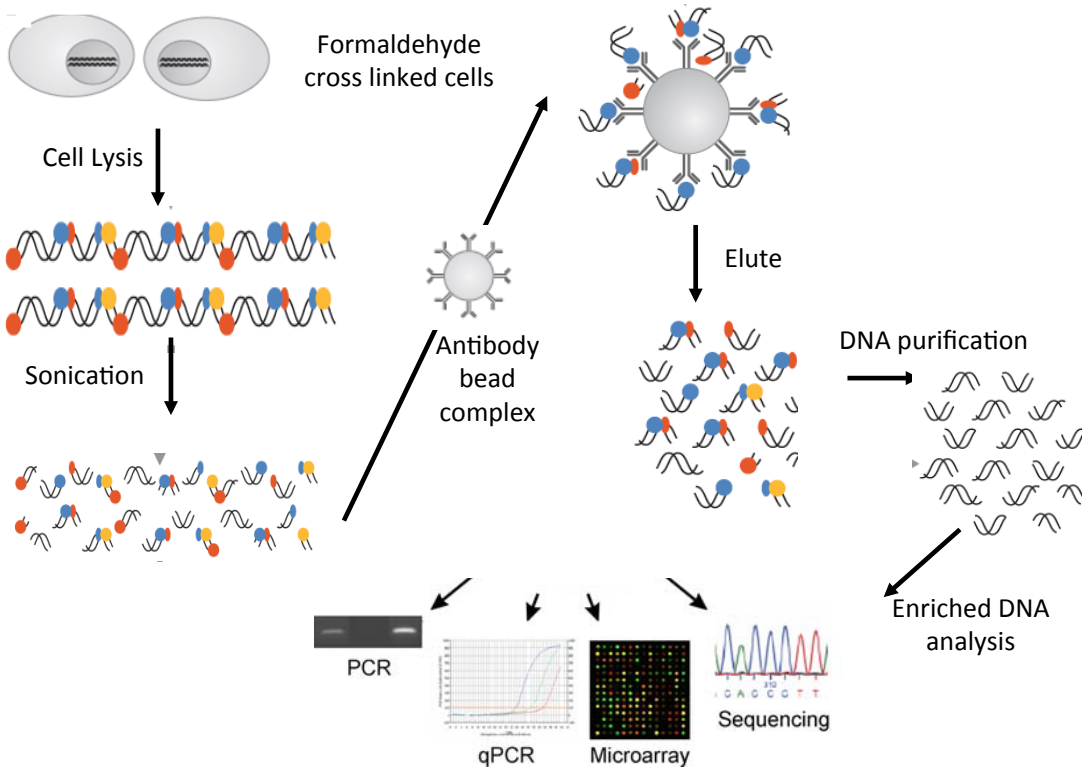


Figure 6. Work flow of in vivo ChIP experiment

The first step in ChIP is establishing crosslinking of the protein and DNA complex within a cell, using a crosslinking agent such as formaldehyde. After crosslinking, cells are lysed and the nuclei isolated, chromatin is released from nuclei by sonication, which fragments the chromatin into shorter pieces. Subsequently, the sheared protein-bound DNA is immunoprecipitated by use of a specific antibody against the protein of interest. Next, protein-DNA complexes are reverse crosslinked to release the enriched DNA. The DNA is then purified, and the resulting abundance of bound DNA is analyzed by different methods, including qPCR, PCR, microarray, and high-throughput sequencing.
(This figure is adapted from Collas, 2010, with modifications).

2.2.4.2 Protein Binding Microarray (PBM)

PBM has been widely used in the study of transcription-factor binding sites within the genome (Bulyk, 2004; Mukherjee et al., 2004). In a typical PBM experiment, the transcription factor is fused to a protein that can be easily detected through the use of an antibody-bound fluorophore, followed by hybridization to an oligonucleotide microarray (Berger et al., 2006; Philippakis et al., 2008). Similar to other microarray-associated techniques, PBM is burdened with disadvantages such as

low resolution and high rates of false positive results.

2.2.4.3 Microcalorimetry

In microcalorimetry techniques, the calorimetry of a small sample is measured. The principle of the technique is based on changes in heat energy that occur during any chemical or physical development. Differential scanning calorimetry (DSC) and isothermal titration calorimetry (ITC) are the two most common microcalorimetric techniques used to study protein-DNA interaction. The DSC technique measures the differences of heat energy changes between samples and reference cells. This is made possible due to the nature of the processes that occur during protein-DNA interactions; during processes such as protein unfolding and complex formation, the thermal stability of proteins change, resulting in the production of heat (Lopez and Makhatadze, 2002). A sample and reference with matched calorimetric cells are enclosed in a chamber, and a computer system is used to measure temperature differences between cells.

The ITC technique is usually applied for direct measurement of DNA and protein interaction/binding. It measures the extent of binding affinity as well as enthalpy and entropy changes (Leavitt and Freire, 2001). In a typical ITC experiment, a solution of either protein or DNA is titrated into a solution of its binding partner. Upon their interaction, the released heat is monitored over time. As the enthalpy of binding is dependent on temperature, this measurement is used to calculate the heat capacity of binding (Lopez and Makhatadze, 2002). The advantages of the ITC technique are its high sensitivity and its ability to directly measure binding constants in the nanomolar to picomolar range (Jelesarov and Bosshard, 1999).

2.2.4.4 Other in vitro techniques

Other in vitro techniques, such Electrophoretic Mobility Shift Assay (EMSA) and DNase I Footprinting (Fox et al., 1999; Hampshire et al., 2007), have been used to study protein-DNA interactions. The DNase I Footprinting assay involves amplification of a fragment of DNA containing the potential binding site by PCR through the use of radioactive or fluorescent-labeled nucleotides. Next, the protein of interest whose binding site is to be identified is bound to the labeled DNA fragment. A cleavage agent, such as DNase I enzyme, is then applied to digest unbound DNA fragments, leaving the binding sites unaffected (Dey et al., 2012).

The EMSA technique has been widely used to study the transcription factors of binding sites due to its high specificity (Garner and Revzin, 1981; Garner and Revzin, 1986; Hellman and Fried, 2007). In a typical EMSA experiment, the linear DNA fragment that contains the potential binding site(s) is used. First, the transcription factor of interest is mixed with radio- or fluorophore-labeled DNA probes. The mixture then is electrophoresed, followed by detection of the separation of the fragments. The transcription-factor-bound DNA complex moves slowly through the gel, while free DNA samples move faster. One of the main drawbacks of the EMSA method is that it does not provide precise direct information on the exact binding site sequences involved (Hellman and Fried, 2007).

2.2.5 Metal interaction with DNA and protein

Metal ions in living cells play important roles in physiological and structural functions. For example, Zn^{2+} acts as co-factor in many enzymatic activities, while divalent cations are essential for transcription, translation, and replication of genetic codes (Izatt et al., 1971). Most importantly, metal ions play a significant role in DNA

structure stabilization through the coordination of positive-charged metal ions that interact with the negative-charged phosphate backbone of DNA (Neidle, 2008). When metal cations are absent, the double helix conformation of DNA is not formed, and DNA is not functional (Sissoeff et al., 1976). On the other hand, some cations can change DNA structure and cause bending in the DNA helix. Additionally, when metal is conjugated with a base pairing between DNA strands, different structural patterns are formed, for example, G-quadruplex or helical junctions (Müller, 2010).

Metalloproteins play very significant roles in many biological processes. Metal ions can either use the information from peptide sequences and contribute to protein folding, or bind to folded proteins to make them functional and enhance their properties. (Dudev and Lim, 2008; Szarka and Lőrincz, 2014). Particularly, copper (Markossian and Kurganov, 2003) and zinc (Cox and McLendon, 2000) metalloproteins play very important roles in all living systems. For instance, zinc-binding domains are found in many gene regulatory proteins associated with nucleic acid binding and gene regulation mechanisms (Laity et al., 2001). Metal binding studies have showed that metal-binding specificity is regulated by the combination of soft/hard acid/base properties and also stabilization of ligand electronic energy effects (Laity et al., 2001).

More than 30% of proteins need the help of co-factors to perform their correct function (Gray, 2003; Wittung-Stafshede, 2002). As proteins are flexible, they may make an imperfect choice for binding to metals. To overcome this problem, affinities for metals follow a universal mode of preference (Waldron and Robinson, 2009). The availability of metals within cells is very limited, causing proteins to compete with other molecules for binding to metals. There are numerous chemical and physical factors within both protein and metal that affect metal preferences, such as ligand

number, characteristics, net charge, polarizability, ionic radius, and coordination geometry (Dudev and Lim, 2014). Zinc and magnesium metals are very popular metal co-factors; as such, competition occurs in cells between enzymes for interaction with zinc and magnesium (Foster et al., 2014; Waldron et al., 2009).

In metal interactions with proteins, any amino acid with electronegative elements in its side chain is able to participate in the interaction (Xiao and Wedd, 2010). Imidazoles of histidines, thiolates of cysteines, water molecules, and carboxylates of glutamic and aspartic acids are considered to be typical ligands for metal binding. As some of these ligands are sensitive to acid/base environments, the release of protons occurs upon metal binding, and consequently, deprotonation results in the forces necessary for successful folding (Blasie and Berg, 2002; Li et al., 2008).

It is a very well-known fact that metal ions accept negative charges from ligands, since they are good electron acceptors. The strength of negative charges from ligands contributes to metal geometry and selectivity (Sakharov and Lim, 2005). Moreover, metal ions with high charges cause polarization of the protein groups surrounding the metal ion. For example, imidazole rings of histidine with delocalized electrons require that metal ions be polarized in order to increase their binding affinity (Sakharov and Lim, 2005; Devereux et al., 2014; Wu et al., 2011).

There are two types of interactions that occur between DNA and metal ions. One is a ligand-mediated interaction, and the other a direct binding interaction between DNA and the metal ion. Direct binding is characterized as an interaction between an empty orbital of the metal ion and the filled orbital of the ligand atom of a nucleobase (Shamsi and Kraatz, 2013). Conversely, ligand-mediated interactions are characterized as interactions that occur between the H-bond and π- π interactions of a ligand of a metal complex containing heterocyclic nucleobases. Such interactions are

facilitated by van der Waals forces, and occur through intercalation or shape-selective binding to the DNA major grove (Barton and Lolis, 1985; Shamsi and Kraatz, 2013). In a study, experimental results demonstrated that the N7 sites of purines constitute the best interaction sites for metal binding; as they are exposed in all DNA forms, such interactions result in base-pairing strengthening (Sponer et al., 1998). Interactions between divalent metal ions (e.g., Zn^{2+}, Ni^{2+}, and Co^{2+}) and B-DNA are an example of direct interactions between DNA and metal ions (Aich et al, 1999; Lee et al., 1993).

2.2.6 DNA topology and secondary structure

Francis Crick and James Watson first proposed the structure of DNA based on X-ray images taken by Rosalind Franklin. A DNA molecule consists of two complementary polynucleotide chains that form a double helix. Each nucleotide consists of a phosphate group, a (2'-deoxyribose) sugar, and one of four bases: Adenine, Thymine, Guanine, or Cytosine. This is typically called B-DNA, characterized as a right-handed helix with 10.5 base pairs per turn at physiological conditions. The topological characteristics of DNA are due to the two strands of DNA being multiply intertwined (Mirkin, 2001).

Circular DNA, found in bacteria, provides a very well-established example of topological domain. In circular DNA, there are no DNA ends, as both strands of DNA are covalently closed. The supercoiling DNA is similar to a normal DNA helix, coiled into a helix of a higher order. Local changes in the secondary structure of DNA, which lead to unwinding of DNA, are energetically favorable in negatively supercoiled DNA (Mirkin, 2001).

B-DNA accounts for most of the DNA behavior that occurs within the cell. However, DNA is not always present in this form, and can form other structures,

including cruciform and Z-DNA, among others (Zhao et al., 2010). Additionally, hairpin structures can be formed by DNA sequences with inverted repeats (IRS), or single stranded DNA (ssDNA). Proteins associated with transcription recognize such secondary structures (Zhao et al., 2010). Hairpins are also formed from double stranded DNA (dsDNA), including cruciforms with two opposite hairpins that extrude by intra-strand base pairings from a palindromic sequence (Panayotatos and Wells, 1981; Platt, 1955).

Cruciforms were first recognized soon after Watson and Crick's discovery (Panayotatos and Wells, 1981; Platt, 1955). The negative supercoiling of dsDNA free-energy results in cruciform stabilization. Three different ways have been reported describing how hairpins interact with proteins involved in the biological processes of cells. One is through cruciform formation that changes the DNA coiling state, which is believed to influence the binding of regulatory proteins associated with transcription and recombination (Cozzarelli and Wang, 1990; de la Cruz et al., 2010; White and Bauer, 1987). As well, hairpins have been reported to be able to inhibit interactions between DNA and protein if they obstruct the recognition site of the protein (Horwitz and Loeb, 1988). There are numerous reports on direct bindings of protein to DNA hairpins found in the literature as well (Barabas et al., 2008; Gonzalez-Perez et al., 2007; MacDonald et al., 2006; Masai and Arai, 1997; Val et al., 2005).

Two main mechanisms have been described that allow for the formation of cruciforms. The first type, referred to as type S, suggests small-scale melting of the DNA double helix at around IRs. Such a small opening facilitates the pairing of a few bases, with their related bases in the repeat. The stem can then be elongated further through the help of negative supercoiling. The other type, referred to as type C, is

characterized by the melting of a large region of DNA, and is favoured by close AT-rich sequences. This eventually allows the hairpins to fold on both strands and form a cruciform. Type S is highly dependent on ionic conditions and IRs (Sullivan and Lilley, 1987), while type C only occurs at low-salt conditions and in the presence of AT-rich sequences, which are critical to this type of process (Bikard et al., 2010; Murchie and Lilley, 1987.)

There are many regions within genomic DNA that have the potential to form hairpins or cruciforms. However, most of these regions are unlikely to contribute to the formation. To illustrate, plasmids containing the DNA of *Drosophila* around the regions surrounding heat shock genes *hsp70, hsp22*, and *hsp26* have been shown to display sensitivity to S1 nuclease enzyme (Mace et al., 1983). Nevertheless, these regions are not involved with IRs; instead, hairpins are more likely formed in regions with high purine content. Additionally, when integration host factor (IHF) binds to an upstream activator region of promoters in the plasmid DNA of *E. coli,* the sensitivity of the DNA found in the promoter site to chemical modification by KMnO4 (a probe for ssDNA) is enhanced. Such structural modifications thereby increase the transcription rates in the promoter (Parekh and Hatfield, 1996).

Many examples detailing the importance of hairpins or cruciforms in transcriptions have been reported involving eukaryotes. For example, high mobility group (HMG) chromosomal DNA binding proteins have been shown to bind DAN secondary structures. These proteins are both dsDNA and ssDNA binding proteins, and similar to histones, are involved with chromosomal function, structure, and transcription. Indeed, HMG proteins have been observed activate DNase I-sensitive sites in chicken erythroblast extracts chromatin (Aravind and Landsman, 1998). Similarly, the pig thymus HMG-1 DNA binding protein in pBR322 plasmid was

demonstrated to bind to regions with IRs, resulting in the removal of a cruciform-structured transcriptional block from DNA (Einck and Bustin, 1985; Waga et al., 1990). In other work, it has been shown that cruciform structures bind to the regulatory domain II of the HMG-I protein responsible for enhancing transcriptional activity of this gene (Robbe et al., 1998). In the same way, DNA-binding proteins that can bind to the secondary structures of DNA to activate transcription, such as hairpins and cruciforms, can also block transcription as well by overlaying or interfering with the recognition sites of DNA-binding proteins (Wadkins, 2000).

Non-transcription factors also interact with hairpin or looped ssDNA. Under proper salt conditions, ssDNA with guanine repeats can form DNA quadraplexes (Williamson, 1994). They are formed at the telomeres of chromosomes, and are recognized by telomere-binding proteins (Laporte et al., 1998). Hairpins are also associated with recognition of the quadraplexes by telomere-binding proteins. The telomere-binding protein has been shown to first bind to a hairpin, which results in disruption of G/G base pairs within the hairpin; consequently, the fashioned ssDNAs join to form a quadraplex. Such binding stabilizes the quadraplex structure and prevents a telomerase binding from occurring (Wadkins, 2000).

CHAPTER 3 – MATERIALS AND METHODS

3.1 Chemicals

All consumable chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise stated.

3. 2 Bacterial media

Liquid and solid media (2YT and/or Luria-Bertani (LB)) were prepared and autoclaved before use. Stock solutions of Kanamycin (Km) antibiotic (Km50 = 50mg/ ml) were prepared in water (filter sterilized) and stored at 20°C.

3.3 Antibodies

Anti-His and anti-Strep antibodies were purchased from GenScript (Piscataway, NJ). Anti-Strep antibody was raised in rabbits against the immunizing peptide "NWSHPQFEK". Anti-His antibody was raised in mouse against the immunizing peptide "HHHHHH". Fluorophore-coupled goat anti-mouse and goat anti-rabbit antibodies were purchased from Mandel (Guelph, ON).

3.4 Plasmid transformation: Electroporation of electro-competent bacterial cells

25µl of electro-competent cells were added to 5µl of plasmid DNA and 20µl of MQ-H_2O, for a total volume of 50µl. The obtained cell solution was then transferred to a cooled 0.2 cm gap electroporation cuvette. Cells were electroporated with a Gene Pulser (Bio-rad, Mississauga, ON) set to capacitance 25µFD, resistance 200 ohms, and voltage 2.5 kvolts. Cells were then added to 500 µl of 2YT with 1M glucose solution, and incubated at 37°C for 1hr, with agitation set to 250 rpm. Next,

cells were plated on LB agar selective plates using Km50 antibiotic.

3.5 Protein expression in *E. coli* cells

The *E. coli* BL21 derivative strain DE3 (+) was used for expression of TGA2, TGA2-VFI, NPR1 Full-length, NPR1-Δ374, NPR1-Δ513, NPR1- Δ535, NPR1-Δ463, NPR1-Δ374-570, NPR1-Δ374-580, NPR1-Δ374-535, NPR1-Δ463-570, NPR1-Δ463-580, NPR1-Δ463-535, NPR1 full-length-580, NPR1 full-length-570, NPR1 full-length-535, NPR1-POZ, NPR1-Δ110, NPR2, NPR3, NPR4, NPR5, NPR6, *Zm*NPR1, and *Os*NPR1 proteins. All genes were amplified by PCR cloned to create N-terminal fusions with a 6X His tag or Strep tag. 2 quantities of 5mL of an overnight culture, grown at 37°C with agitation set at 250rpm, were used to inoculate 2 quantities of 500mL cultures for *E. coli* cells producing the abovementioned proteins. All cultures were grown at 37°C with agitation set at 250 rpm until they reached an OD600 of 0.4-0.6, at which point they were induced with 1mM IPTG (Isopropyl β-D-1-thiogalactopyranoside; BioShop, Mississauga, ON) and grown for an additional 2.5 to 4 hours at 37°C and agitation set at 250rpm. Bacterial cells were collected by centrifugation, set at 4°C and 8000 rpm for 5 minutes (Sorvall-RC-5C-Plus; Mandel Scientific Company, ON). Pellets then re-suspended in 30ml of distilled cold MQ-water and centrifuged once more at 3500 rcf for 15 min at 4°C. The SN was then removed and cells stored at -20°C.

3.6 Protein extraction / Cell lysis

Approximately 2g of cell pellets were used in each replicate of cell lysis. The pellets were re-suspended in 16mL of filtered IP buffer (0.1 % SDS, 1 % Triton X-100, 50mM HEPES pH7.9, 150mM NaCl, 1mM EDTA) and lysated using the

EmulsiFlex C3 (AVESTIN, CA) for 2 minutes, with cooling set to 24°C. All samples were centrifuged at maximum speed at 4°C for 30 minutes. Supernatants (SN) were then collected and immediately used for modified in vitro ChIP experiments.

3.7 Western blot analysis

Prior to bulk protein expression for in vitro ChIP experiments, solubility and expression of fusion proteins produced in *E. coli* was confirmed by performing immunoblot analysis. Anti-His or/and anti-Strep antibodies were used as primary antibodies (GenScript, Piscataway, NJ). Fluorophore coupled goat anti-mouse and goat anti-rabbit antibodies were used as secondary antibodies (Mandel, Guelph, ON). First cleared SN was removed from the pellet followed by acetone-precipitation, and washing with binding buffer (20mM Tris pH 7.9, and 50mM NaCl). Next, protein samples were eluted with 30µL of 1X sample buffer (0.06M Tris-Cl pH 6.8, 2% SDS, 1% glycerol, 0.025% bromophenol blue) and heated at 100°C for 5 minutes. Protein samples were pelleted by centrifugation (max speed, 30 seconds), followed by SN collection and electrophoresis on SDS-polyacrylamide gel (12% and/or 10%). Next, gels were transferred to nitrocellulose, and reacted according to the manufacturer's instructions (LI-COR, Lincoln, NE) with anti-Strep antibodies, followed by detection with Goat anti-Rabbit IRDye® 680LT. His-tagged proteins were reacted with anti-his antibody, followed by detection with Goat anti-Mouse IRDye® 800LT.

3.8 Modified in vitro ChIP method for the study of NPR1 binding to PR1 promoter

In vivo ChIP method has been very well-established in the study of DNA-protein interactions in *Arabidopsis* plants (Chakravarthy et al., 2003). For the

purposes of this study, the in vivo method was modified so that in place of plant tissues, *E. coli* recombinant proteins and DNA sequences cloned into plasmids could be used as samples. Plasmid DNA contained different PR1 promoter sequence constructs. 1mL of SN after cell lysis (protein source of the assay) was mixed with 1µl of plasmid DNA (DNA source of the assay). Plasmid DNA concentrations were normalized through several dilutions to concentrations that yielded approximate CT values of 20 in qPCR assays. Analyses by qPCR were performed using specific primers on the PR1 promoter (Table S3). 100 µl of 10X protease inhibitor cocktail mixture (Sigma-Aldrich) was added to the mixture of protein and DNA, followed by incubation at room temperature for 15 min using Roto-Torque with slow spin. Crosslinking of protein and DNA was accomplished by addition of 28 µl of formaldehyde from 0.01% stock solution. The obtained mixture was incubated at room temperature for 5 min using Roto-Torque with slow spin. To extinguish the crosslinking process, 54µl of glycine from 125mM stock solution was added to the mixture. Subsequently, the mixture was incubated at room temperature for 5 min using Roto-Torque with slow spin. After crosslinking the protein and DNA of interest, 1µl of genomic DNA from *Arabidospsis* wild-type leaf tissue was added to the mixture and incubated for 5 min at room temperature using Roto-Torque and spin slowly; this step was taken to normalize the *PR1* enriched DNA from *PR1* gene, as the *Ubiquitin5* gene present in genomic DNA is constitutively expressed (Callis et al., 1990). Similar to plasmid DNA concentrations, genomic DNA concentrations were also diluted several times, followed by qPCR with *Ubiquitin5* gene-specific primers, to reach approximate CT values of 20. Next, a protein-antibody bound bead complex was prepared using 1:200 dilutions of anti-Strep and anti-His antibodies obtained from stock with 0.5mg/ml concentration. 50µl of protein A-agarose support beads

were spun down; the SN was then discarded and the beads (Bio-Rad, Mississauga, ON) washed 3 times with 500µl of IP buffer (0.1 % SDS, 1 % Triton X-100, 50mM HEPES pH7.9, 150mM NaCl, 1mM EDTA). Each washing included spinning down the beads, removing the cleared SN and gently re-suspending the beads in IP buffer. After washing steps beads were then re-suspended in 500µl of IP buffer, followed by addition of antibody (either anti-Strep or anti-His) to the mixture and incubation at room temperature for 1-1.5hrs, using Roto-Torque with slow spin. After 1-1.5hrs antibody-bound beads complex was washed 2 times with IP buffer. The protein-bound DNA complex samples were immunoprecipitated by addition of the beads-bound antibody mixture, followed by incubation overnight at 4°C using Roto-Torque with slow spin. Following overnight incubation, the protein bound DNA complex was reverse cross-linked to release the enriched DNA (DNA bound to protein). The beads were washed 3 times with IP buffer and 2 times with 1X TE buffer (100mM Tris and 50mM EDTA). Beads were then re-suspended in 200µl 1X TE +1% SDS and incubated at 65°C using a pre-heated water bath for 15 min. Beads were then spun down, and SN collected. 400µl of 1X TE +1% SDS buffer was added to the mixture, which was incubated overnight at 65°C using a water bath. Following overnight incubation, the DNA was precipitated and subjected to qPCR with 40 cycles of two-temperature protocol, for a total volume of 20µl, using the Bio-Rad (Mississauga, ON) iQ SYBR Green Supermix kit according to the manufacturer's instructions. For all studies, 3 biological replicates of each sample were used. Recorded CT values were compared to those of *UBQ5* primer pair products; final values were $2^{(CTUBQ5-CTPR1)}$. The PCR primer pairs specific to the PR1 promoter and UBQ5 are listed in Table S3.

3.9 Chemical addition during modified in vitro ChIP experiments

During modified in vitro ChIP experiments, SA, INA, 4-OH-BA, Isotianil, and EDTA (as metal chelator) were added to test their effect on the binding of NPR1 to the PR1 promoter. Different concentrations of each chemical were tested, including 0.1μM, 1μM, 10μM, 80μM, and 10mM (just for EDTA). Chemical solutions were prepared using MQ-H_2O with stirring for approximately 20 min at room temperature. For preparation of INA solution, INA was first dissolved in 100μl of Dimethyl Sulfoxide (DMSO), then added slowly to MQ-H_2O with stirring. Chemical solutions were added to protein and DNA mixtures prior to cross-linking step. Mixtures were first incubated for 15 min at room temperature using Roto-Torque with slow spinning, then submitted to cross-linking.

3.10 Denaturation of double stranded plasmid DNA

The heat denaturation method was applied towards the denaturation of double-stranded plasmid DNA to single-stranded form. The method was simple and efficient. 1μM of plasmid DNA was heated to approximately 100 °C, boiled for 2 min, then left to stand for 2 minutes in an ice bath. The resulting denatured plasmid was then incubated with protein, and subsequently submitted to crosslinking.

3.11 Plasmid restriction enzyme digestion

Two different restriction enzymes, *SnaBI* and *BglII* (New England BioLabs, Beverly, MA), were used to digest the prepared plasmid fragments of the PR1 promoter (-701 and -870 fragments). The reaction mixture contained 10 units of each enzyme, as per the manufacturer's instructions, 2 μl of the manufacturer-recommended reaction buffer (Buffer 3 for *BglII* and buffer B for *SnaBI*), 1μM of

plasmid DNA, and sterile filtered water, for a final volume of 20 µl. The reaction mixture was incubated at 37°C for 2 hr. Next, 2µl of stop buffer (water, 50% glycerol, 100 mM EDTA, 1% sodium dodedyl sulphate (SDS), 0.1% bromophenol blue) was added to extinguish the reaction. The digested plasmid PR1 promoter (Figure S9) fragments were incubated with protein, and then submitted to cross-linking.

3.12 4-OH-BA addition as DNA inhibitor during modified in vitro ChIP

Four different 4-OH-BA concentrations were selected to test the effect of 4-OH-BA on the binding of NPR1 to the PR1 promoter. SA and 4-OH-BA solutions were made as described in section 3.11. 100 µM from different stock solutions (Table 3.1) were added to a protein-plasmid DNA mixture, followed by incubation for 15 min at room temperature using Roto-Torque with slow spin. Next, 100 µM of SA from different stock solutions (Table 3.1) were added to the mixture, followed by incubation for another 15min at room temperature using Roto-Torque with slow spinning. Different stock solutions of 4-OH-BA were added to each biological replicate, while concentrations of SA were kept constant (Table 3.1). After addition of 4-OH-BA and SA, samples were submitted to cross-linking. For modified in vitro ChIP experiments with NPR1-Δ513 and NPR1-Δ533 (Section 4.10 Figure 14), 4-OH-BA concentrations (1, 0.01, 0.00001, 0.000001, and 0.0000001µM) were added individually to protein and DNA mixtures, then incubated for 15 min at room temperature using Roto-Torque with slow spinning, followed by a crosslinking step.

Table 3.1. Combinations of SA and 4-OH-BA concentrations used for modified in vitro ChIP analysis

Chemical names and Stock concentrations	
1μM SA	10μM SA
1μM SA + 1μM 4-OH-BA	10μM SA + 1μM 4-OH-BA
1μM SA + 10μM 4-OH-BA	10μM SA + 10μM 4-OH-BA
1μM SA + 80μM 4-OH-BA	10μM SA + 80μM 4-OH-BA

3.13 Plant Growth Conditions and Transformation

Conditions for growth of Arabidopsis thaliana (Columbia) and *PAL* mutant plants were as previously described in the literature (Liu et al., 1995).

3.14 Treatment of *Arabidopsis thaliana* plants with SA, 4OH-BA, 3OH-BA, BA, INA, and catechol for qRT-PCR experiments

Arabidopsis (ecotype Columbia, and *PAL* single mutant) leaves from 6-week-old plants grown at 21 °C (day) and 18 °C (night) with a 10 hour photoperiod were sprayed with 300μM SA and combinations of SA and 4-OH-BA, 3-OH-BA, BA, INA, and catechol (Table 3.2). For INA analysis, INA was first dissolved in 100μM of DMSO, added to Milli-Q water slowly by stirring, then used to spray leaves. Leaf tissues were harvested at specific time periods: prior to treatment and 12hrs after treatment.

Table 3.2. Combinations of SA with INA, 4-OH-BA, 3-OH-BA, BA, and Catechol concentrations used for RT-PCR analysis

Chemical names and Stock concentrations				
300µM SA	300µM SA	300µM SA	300µM SA	300µM SA
300µM SA + 30µM INA	300µM SA + 30µM 4-OH-BA	300µM SA + 30µM 3-OH-BA	300µM SA + 30µM BA	300µM SA + 30µM Catechol
300µM SA + 3µM INA	300µM SA + 3µM 4-OH-BA	300µM SA + 3µM 3-OH-BA	300µM SA + 3µM BA	300µM SA + 3µM Catechol
300µM SA + 0.3µM INA	300µM SA + 0.3µM 4-OH-BA	300µM SA + 0.3µM 3-OH-BA	300µM SA + 0.3µM BA	300µM SA + 0.3µM Catechol

3.15 Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)

Total RNA was extracted from leaves using the RNeasy plant mini-kit (Qiagen, Mississauga, ON) in accordance with the kit instructions. RNA was first treated with DNase I (Invitrogen, Carlsbad, CA). With the use of SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA), first strand cDNA synthesis was then generated with the (dT) 17VN oligo in the presence of 0.4U 133 RNasin (Fisher Scientific, Pittsburg, PA). The newly synthesized cDNA was diluted to 1/200, for a final concentration of 10 ng/L of input RNA. Quantitative Polymerase Chain Reaction (qPCR) was performed on a C1000^{TM} Thermal Cycler (Bio-Rad, Mississauga, ON) using a two-temperature cycling regimen initiated with 15 min activation at 95°C, followed by 40 cycles, lasting 2 min each, of annealing and extension at 66°C, and 10 sec denaturation at 95°C. Each reaction contained 5 ng cDNA, iQ SYBR Green Supermix kit 1 X SYBR Green® (Bio-Rad, Mississauga, ON) and 0.5pmol oligonucleotides. Fluorescence data was collected at the end of each analyzed PCR cycle using the absolute quantification CT method (Rutledge and Stewart, 2008). Analyses with qPCR were performed using *PR1* and *Ubiquitin5* primer pairs. Since

the *Ubiquitin5* gene is constitutively expressed (Callis et al., 1990), the *PR1* expression levels were normalized to the *Ubiquitin5* transcript. The sequences of the used primers are as follow: $PR1_F$ 5'-GATCACCGATTGACATTGTA-3', $PR1_R$5'-GAACACAAAAGTAGATCGGT-3', $UBQ5_F$ 5'-GACGCTTCATCTCGTCC-3', and $UBQ5_R$5'-GTAAACGTAGGTGAGTCCA-3'.

3.16 DNA extraction

Genomic DNA was extracted from leaves using the DNeasy plant mini-kit (Qiagen, Mississauga, ON) in accordance to the kit instructions. Next, 5µl of DNA sample was loaded on 1% agarose gel in 1x TAE buffer (12 mM Tris, 6 mM sodium acetate, 0.3 mM EDTA, pH 8.0), including 0.5µg ml^{-1} ethidium bromide. The DNA fragments were separated at 90 V for about 45 min with a mini-sub™ DNA cell (BioRad) in 1x TAE buffer. The DNA was visualized using a ChemiDoc™ MP System (Bio-Rad, Mississauga, ON).

3.17 Data analysis for modified in vitro ChIP

Recorded cycle threshold (CT) values for PR1 promoter primers were compared to those of UBQ5 primer pair products. The final values are $2^{(CTUBQ5-CTPR1)}$. The obtained amounts of NPR1 or TGA2 protein bindings to the *PR1* and *ubiquitine 5* genes promoters were quantified as *PR1* and *UBQ5* enrichments. *PR1* enrichment values were then normalized to *UBQ5* enrichment values (fold enrichment method). Final values were considered as relative *PR1* enrichment. An average of three biological replicates were used to plot graphs and conduct two-tailed Student's t test analyses.

3.18 Statistical Methods

All pooled data are expressed as averages; error bars represent standard deviations. Statistical significance was assessed using a two-tailed Student's t test for comparisons of data from two independent populations.

CHAPTER 4 – RESULTS

4.1 Validation of modified in vitro Chromatin Immunoprecipitation (ChIP) method

The primary purpose of this study was to determine whether NPR1 contains a DNA binding domain(s) that allows for its direct interaction with the PR1 promoter. The scientific reasoning behind this line of inquiry is based on the in vivo ChIP results obtained by Rochon et al. (2006), who demonstrated that NPR1 interacts with the PR1 promoter through an unknown protein X. To address the abovementioned inquiry, an in vitro modified version of the Chromatin Immunoprecipitation (ChIP) method was developed.

In order to develop and validate the modified ChIP method, the TGA2 protein was used as a positive control, as it has been well-demonstrated that TGA2 interacts with the PR1 promoter on two different sites, Linker-scanning 5 and 7 (LS5 and LS7) (Figure 1A) (Lebel et al., 1998). As a negative control measure, VFI-TGA2, a shorter form of the TGA2 protein, was selected due to the absence of the necessary domain (amino acid 1-43) for DNA binding (Figure 1A). In addition, a PR1 promoter sequence (-1 to -1023bp) cloned into a plasmid was obtained as a DNA source.

TGA2 and VFI-TGA2 proteins were tagged with the Streptavidin-binding peptide (NWSHPQFEK) and used for the immunoprecipitation step in modified in vitro ChIP experiments. A western blot analysis was conducted using anti-Strep antibody to confirm the solubility of *E. coli* recombinant expressed TGA2 and VFI-TGA2 proteins and the effectiveness of the antibody; results confirmed the right expression of proteins and specificity of the antibody.

Similarly to classical ChIP methodology, an optimization of parameters, namely concentration amounts and incubation time required for cross-linking of protein and DNA complex, was conducted using formaldehyde as a fixative reagent. Incubation of TGA2, VFI-TGA2, and the plasmid-containing PR1 promoter sequence for 5 min, using a 0.01% formaldehyde concentration, yielded statistically larger relative *PR1* enrichment amounts from the TGA2-PR1 promoter complex as compared to the relative *PR1* enrichment amounts obtained from the VFI-TGA2-PR1 promoter complex (Figure 1B) (Student two-tailed t-test, p-value is 0.0389; p-value < 0.05).

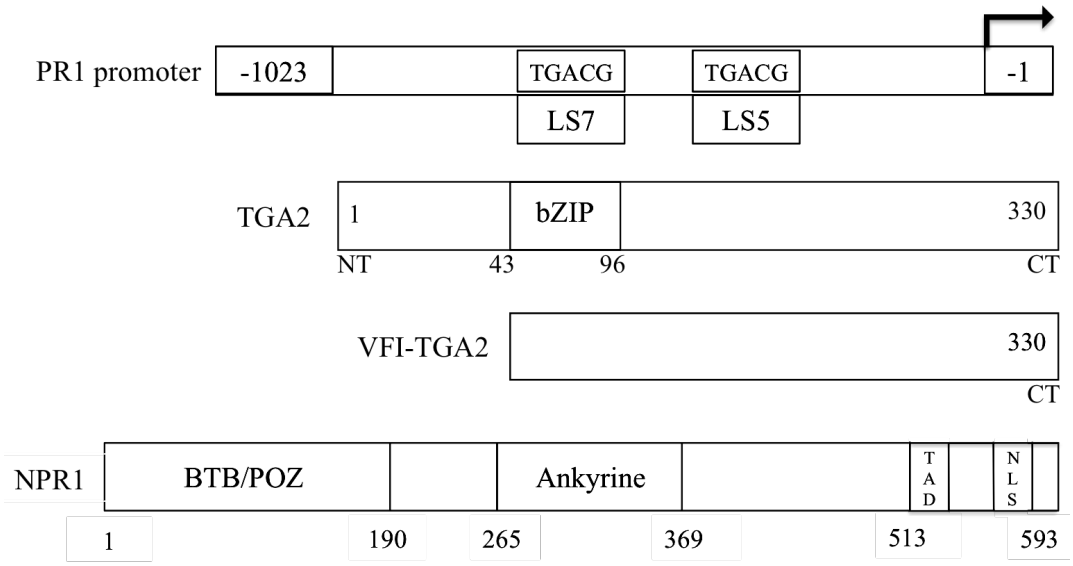
In addition to formaldehyde concentration and cross-linking time, the concentration of antibody required to immunoprecipitate the protein-bound DNA complex was also optimized. For this, a series of anti-Strep antibody dilutions were performed, followed by immunoprecipitation and qPCR analysis. Next, the relative amount of *PR1* enrichment from the TGA2-PR1 promoter complex was compared with values obtained for the VFI-TGA2-PR1 promoter complex. As shown in Figure 1C, upon immunoprecipitation by application of a 1:200 dilution (0.00125mg/ml) of anti-Strep antibody to the protein-bound DNA complex, the relative amount of *PR1* enrichment was found to be significantly larger for the TGA2-PR1 promoter complex than the value obtained for the VFI-TGA2-PR1 promoter complex (Student two-tailed t-test, p-value is 0.0052; p-value < 0.05).

Following validation of the modified in vitro ChIP method using the TGA2 protein as a candidate protein, a similar method was applied to study the binding of NPR1 to the PR1 promoter. Similar to the previously taken approach, TGA2 and VFI-TGA2 were chosen as positive and negative controls, respectively. Results showed that the relative amount of *PR1* enrichment obtained from the NPR1- PR1 promoter

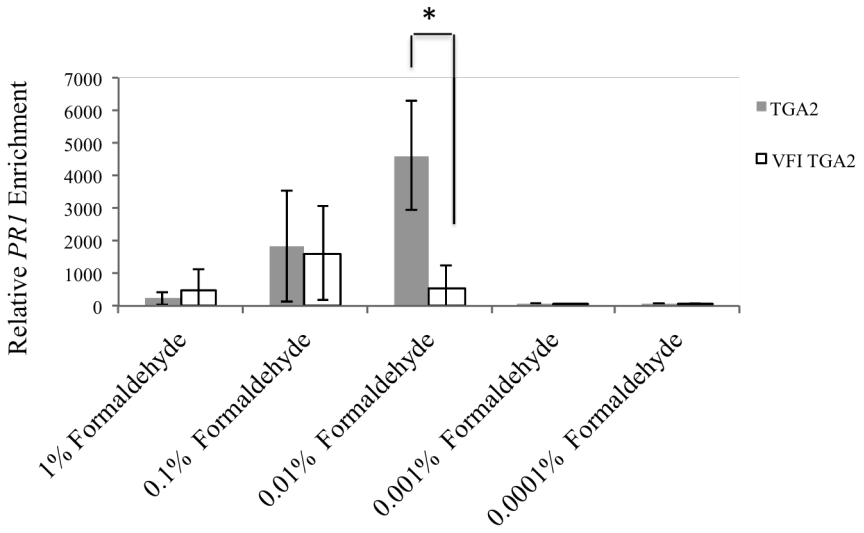
complex was similar to the relative amount of *PR1* enrichment obtained from the TGA2-PR1 promoter complex (Figure 1D) (Student two-tailed t-test, p-value is 0.6008; p-value < 0.05). However, the relative *PR1* enrichment value obtained from the NPR1- PR1 promoter complex was found to be statistically larger than the value obtained from the VFI-TGA2-PR1 promoter complex (Figure 1D) (Student two-tailed t-test, p-value is 0.0460; p-value < 0.05). Similar to previous results (Figure 1B and C), the relative amount of *PR1* enrichment obtained from the TGA2-PR1 promoter complex was statistically larger than the relative amount of *PR1* enrichment obtained from the VFI-TGA2-PR1 promoter complex (Figure 1D) (Student two-tailed t-test, p-value is 0.0111; p-value < 0.05).

These results confirmed the effectiveness of the modified in vitro ChIP method at the optimized parameters of 0.01% concentration of formaldehyde and 1:200 dilution of anti-Strep antibody. In addition, the validated modified in vitro ChIP method was shown to be appropriate to study the binding of NPR1 protein to the PR1 promoter. Finally, the results clearly demonstrated that NPR1 is recruited at the PR1 promoter site in a similar fashion to the TGA2 protein.

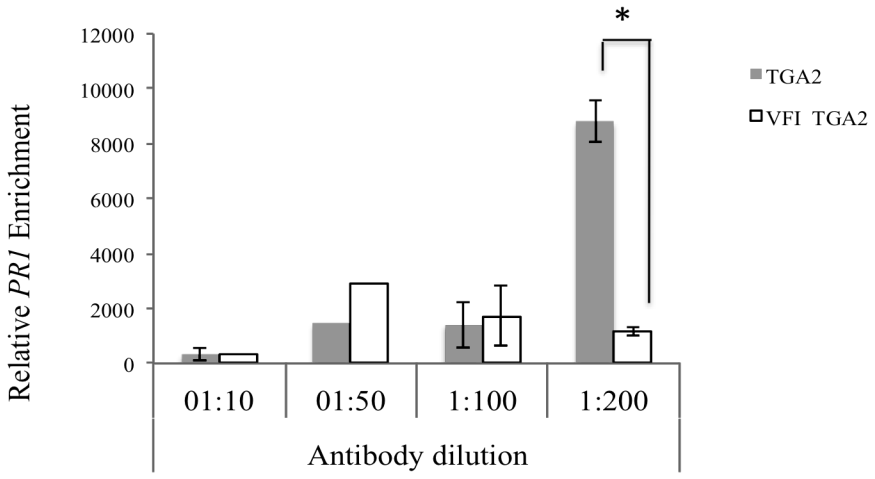
A


PR1 promoter
-1023
TGACG
TGACG
-1
LS7
LS5
TGA2
1
bZIP
330
NT
43
96
CT
VFI-TGA2
330
CT
NPR1
BTB/POZ
Ankyrine
T A D
N L S
1
190
265
369
513
593

B


*
Relative PR1 Enrichment
7000
6000
5000
4000
3000
2000
1000
0
TGA2
VFI TGA2
1% Formaldehyde
0.1% Formaldehyde
0.01% Formaldehyde
0.001% Formaldehyde
0.0001% Formaldehyde

C


*
Relative PR1 Enrichment
12000
10000
8000
6000
4000
2000
0
TGA2
VFI TGA2
01:10
01:50
1:100
1:200
Antibody dilution

D

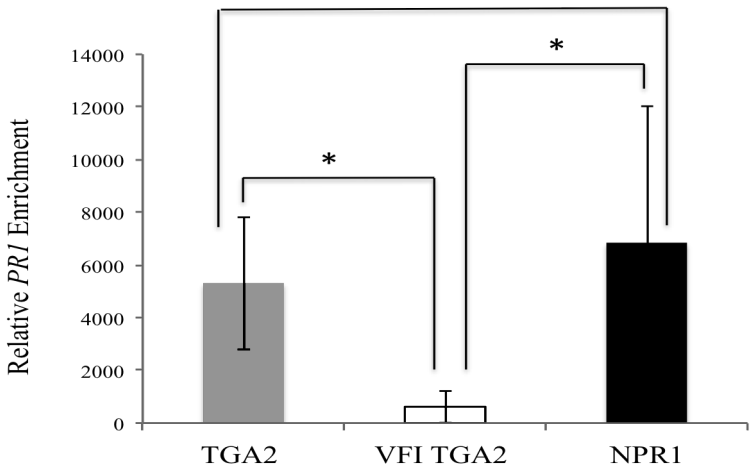


Figure 1.Validation of modified in vitro Chromatin Immunoprecipitation (ChIP) method

A. Schematic representation of protein and DNA constructs used for modified ChIP experiment. Diagram of PR1 promoter indicates the -1 to -1023bp sequence of the promoter used as DNA source in modified ChIP experiments. The TGA2 protein-binding sites on the PR1 promoter are shown in two separate boxes, labelled as LS5 and LS7. The arrow indicates the start of transcription. The full-length TGA2 protein and a shorter version, named VFI-TGA2, were used as positive and negative controls, respectively, in the modified ChIP experiment. Diagram of NPR1-full length protein shows the BTB/POZ domain, Ankyrine repeat, C-terminus transactivation domain, and nuclear localization signal.
B. Histogram illustrating modified ChIP experiments performed on TGA2 and VFI-TGA2 proteins, confirming optimization with 0.01% formaldehyde for cross-linking of the TGA2-bound PR1 promoter complex. ChIPs experiments were performed with anti-Strep antibody as an immunoprecipitator. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).
C. Histogram illustrating modified ChIP experiments performed on TGA2 and VFI-TGA2 proteins, confirming optimized factor of 1:200 dilution of anti-Strep antibody for specific immunoprecipitation of TGA2-bound PR1 promoter complex. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).
D. Histogram illustrating modified ChIP experiments performed on TGA2, VFI-TGA2, and NPR1 proteins, confirming direct binding of NPR1 to the PR1 promoter. TGA2 and VFI-TGA2 used as positive and negative controls, respectively. ChIPs were performed using anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The start (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

4.2 Determination of multiple binding sites between NPR1 and the PR1 promoter

Upon confirmation that the NPR1 protein directly interacts with the PR1 promoter (Figure 1D), we became interested in narrowing down the location of the NPR1 interaction site on the PR1 promoter. For this purpose, modified ChIP experiments were performed to investigate the binding of NPR1 to a series of PR1 promoter deletion constructs, covering different regions of the PR1 promoter sequence (Figure 2A).

Our findings revealed that the relative amounts of *PR1* enrichment for the -701 and -870 fragments were much larger in comparison to the statistically insignificant amount found for the -621 fragment region (base pair -1 to -621) of the PR1 promoter (Figure 2B, blue bar). In addition, a lesser relative amount of *PR1* enrichment was found for the -701 fragment region of the PR1 promoter (Figure 2B, green bar) as compared to the -870 fragment region. Interestingly, the relative amount of *PR1* enrichment obtained from the -870 fragment region was found to be similar to the values obtained for the -1023 fragment region of the PR1 promoter (Figure 2B, white bar**)** (Student two-tailed t-test, p-value is 0.3251; p-value < 0.05). However, the relative *PR1* enrichment values obtained for -701 and -621 fragments were found to be statistically lesser than the value obtained for -1023 fragment of PR1 promoter (Figure 2B, green and blue bars) (Student two-tailed t-test, p-values are 0.0363 0.03332; p-value < 0.05). These results suggest that the NPR1 protein might have multiple binding sites on the PR1 promoter.

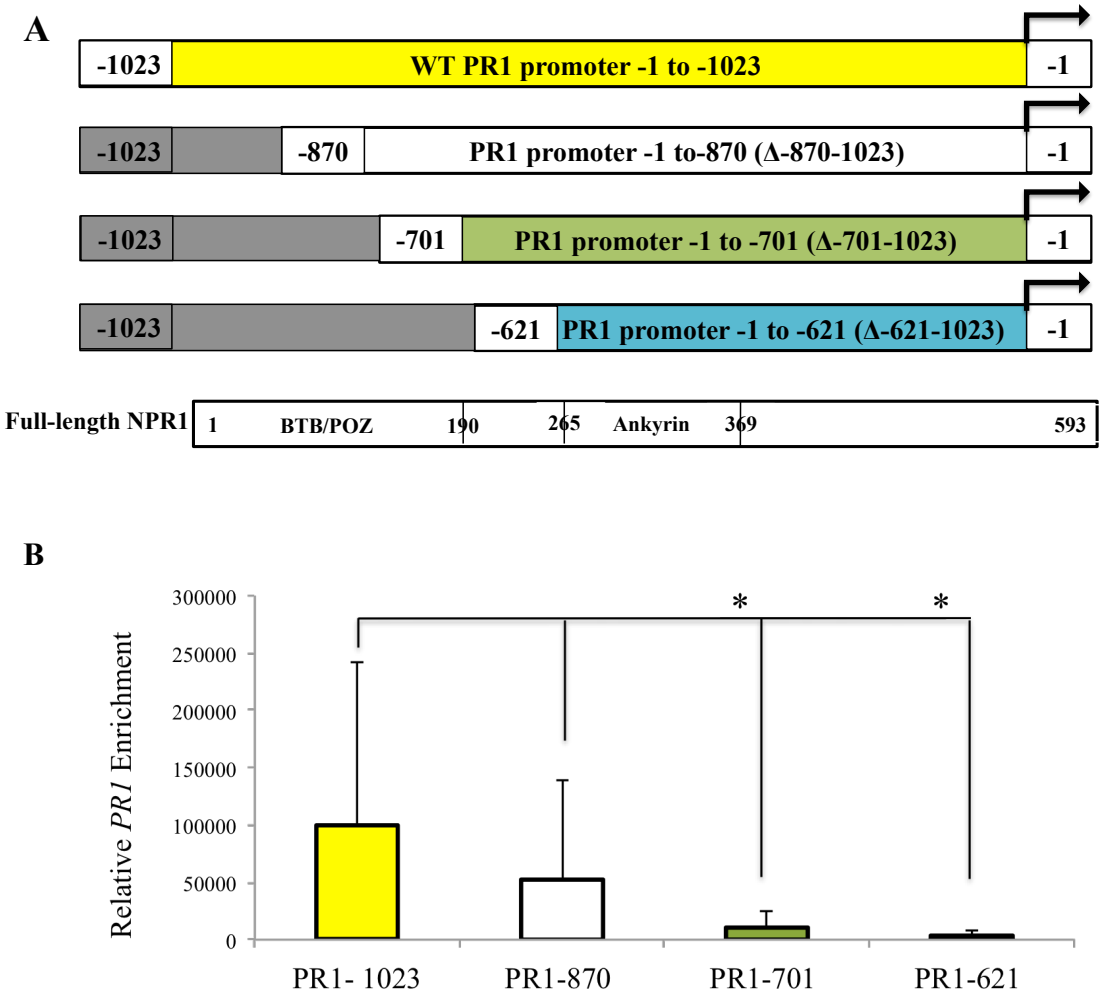


Figure 2. Multiple site binding of NPR1 to the PR1 promoter

A. Schematic representation of PR1 promoter deletion constructs used for modified in vitro ChIP experiments. The arrow indicates the start of transcription. Each box represents a fragment of the PR1 promoter sequence that was cloned into a plasmid and used as DNA source in modified ChIP experiments. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). White box indicates the -1 to -870bp fragment of the PR1 promoter, which sequences from -870bp to -1023bp deleted (Grey colour box). Green box represents the -1 to -701bp fragment of the PR1 promoter that sequences from -701bp to -1023bp deleted (Grey colour box). Blue box illustrates the -1 to -621bp fragment of the PR1 promoter, which sequences from -621bp to -1023bp deleted (Grey colour box). The last white box shows full-length NPR1 protein (amino acids 1-593).

B. Histogram illustrating modified ChIP experiments performed on PR1 promoter deletion constructs, confirming that NPR1 has multiple binding sites on the PR1 promoter. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of more than 6 biological replicates.

4.3 Positive regulation of NPR1 binding to the PR1 promoter by SA and negative regulation by INA and 4OH-BA

In work conducted by Wu et al. (2012), NPR1 was shown to be a specific receptor for salicylic acid (SA). By using equilibrium dialysis, the authors demonstrated that ^{14}C-SA directly binds to NPR1 with a similar affinity as other hormone-receptor bindings.

Since NPR1 is known to be a specific receptor for SA, it remained to be determined whether SA (Figure 3B) affects NPR1 interactions with PR1 promoter. To accomplish this, the cross-linked NPR1-PR1 promoter was incubated with two different concentrations of SA (10 and 80μM). *PR1* enrichment was then determined by modified ChIPs and analyzed by qPCR assays.

Similar to previous results (Figure 2B, white bar), NPR1 was observed to bind to the -870 fragment of the PR1 promoter in the absence of SA. This was determined through the detection of a larger relative *PR1* enrichment before addition of SA followed by gradual decrease in binding relative to incremental SA concentrations (from 10 to 80μM) (Student two-tailed t-test, p-value is 0.0090; p-value < 0.05) (Figure 3C, white bars). In contrast, as can be seen from Figure 3C (blue bars), NPR1 was observed to interact with the -701 fragment of the PR1 promoter in the presence of SA, with significant increases in binding observed in correlation to increments in SA concentration (from 10 to 80μM) (Student two-tailed t-test, p-value is 0.0119; p-value < 0.05). These findings were consistent with previous findings (Figure 2B, green bar); smaller relative *PR1* enrichment amounts were observed for the -701 fragment region promoter in comparison to those obtained for the -870 fragment region. Furthermore, consistent with previous results (Figure 2B, blue bar), *PR1* enrichment within the -621 fragment of the PR1 promoter was not observed either in

the absence or presence of SA (Student two-tailed t-test, p-value is 0.0969; p-value < 0.05) (Figure 3C, blue bars), clearly confirming that the -621 fragment region (base pair -1 to -621) of the PR1 promoter does not interact with the NPR1 protein.

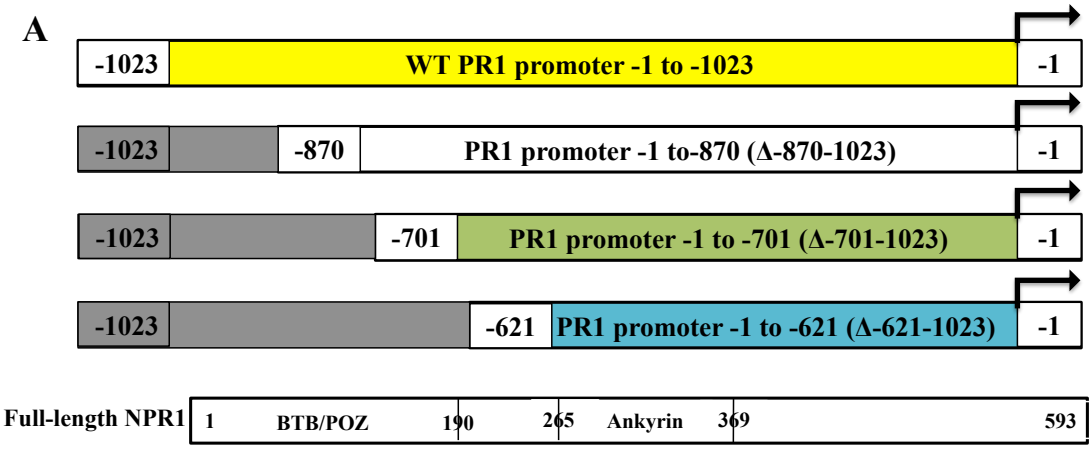
In the aforementioned study, Wu et al. (2012) also demonstrated that NPR1 does not interact with 4-hydroxy benzoic acid (4-OH BA) and 2,6-dichloroisonicotinic acid (INA) (Figure 3B), and that both 4-OH-BA and INA do not trigger *PR-1* expression in *Arabidopsis* leaves. Similarly, for the current work, the effects of both 4-OH-BA and INA (Figure 3D) on NPR1-binding to the PR1 promoter were investigated, using SA as a positive control measure. Similar to our previous approach, modified ChIP experiments were performed with the -701 and -870 fragments of the PR1 promoter; however, the -621 fragment was not further investigated due to previous findings demonstrating that no enrichment occurs for this fragment, neither in the presence nor in the absence of SA (Figure 3C, blue bars). NPR1 cross-linked with either the -701 or -870 fragments were incubated with SA, INA, and 4-OH-BA using two concentrations (0.1 and 1μM), followed by immunoprecipitation using the anti-Step antibody for specific immunoprecipitation of the NPR1-bound PR1 promoter complex.

The obtained results demonstrated that relative *PR1* enrichment was completely abolished for the -701 and -870 fragments of the PR1 promoter in the presence of both concentrations (0.1 and 1μM) of INA and 4-OH-BA (Figure 3D). However, in the presence of SA, the obtained relative *PR1* enrichment amounts for both the -701 and -870 fragments were consistent with our previous results (Figure 2B and 3C, white and green bars). For the -870 fragment region, an increase in relative *PR1* enrichment was observed in the absence of SA, while no enrichment occurred in the presence of SA (Student two-tailed t-test, p-value is 0.0056; p-value <

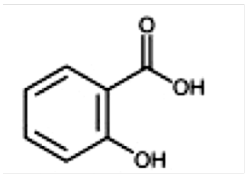
0.05)(Figure 3D, white bars). Similarly, for the -701 fragment, a higher relative *PR1* enrichment amount was observed in the presence of SA, while a decrease in enrichment amounts were noted in the absence of SA (Student two-tailed t-test, p-value is 0.0370; p-value < 0.05) (Figure 3D, green bars).

In conclusion, the obtained results further confirmed that NPR1 directly interacts with the PR1 promoter, and that SA affects NPR1 binding to the -701 and -870 fragments of the PR1 promoter in positive and negative ways, respectively. Moreover, INA and 4-OH-BA were shown to negatively affect NPR1 binding to the PR1 promoter.

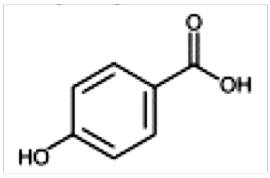
A



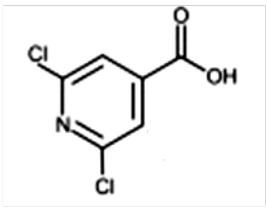
B



Salicylic acid (SA)

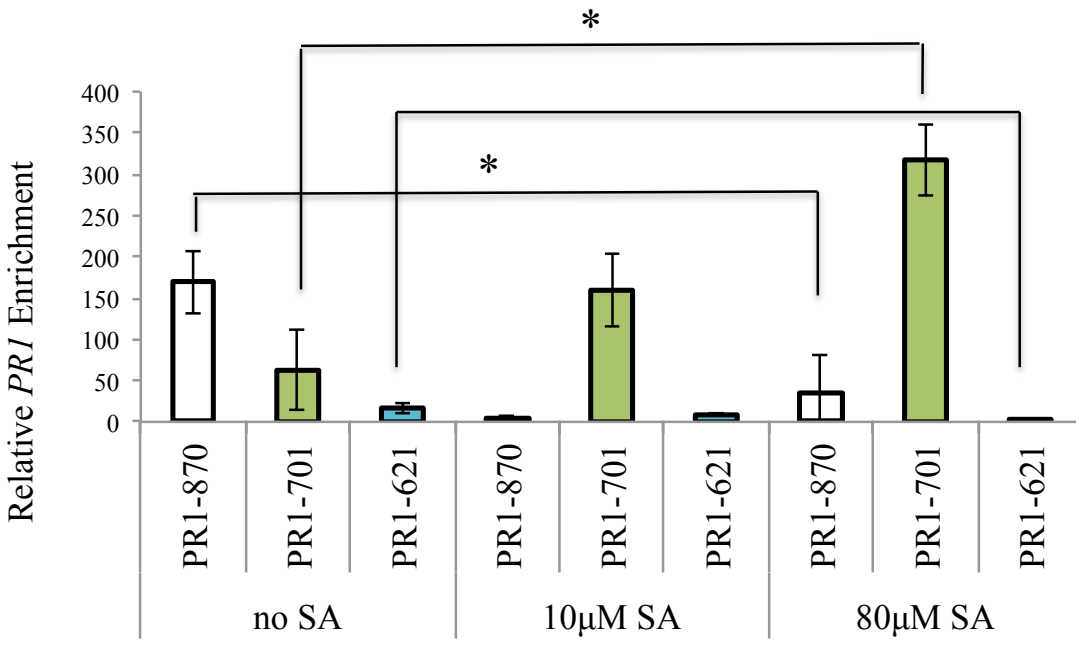


4 Hydroxy Benzoic acid (4-OH-BA)



2,6 Dichloro INA (INA)

C



D

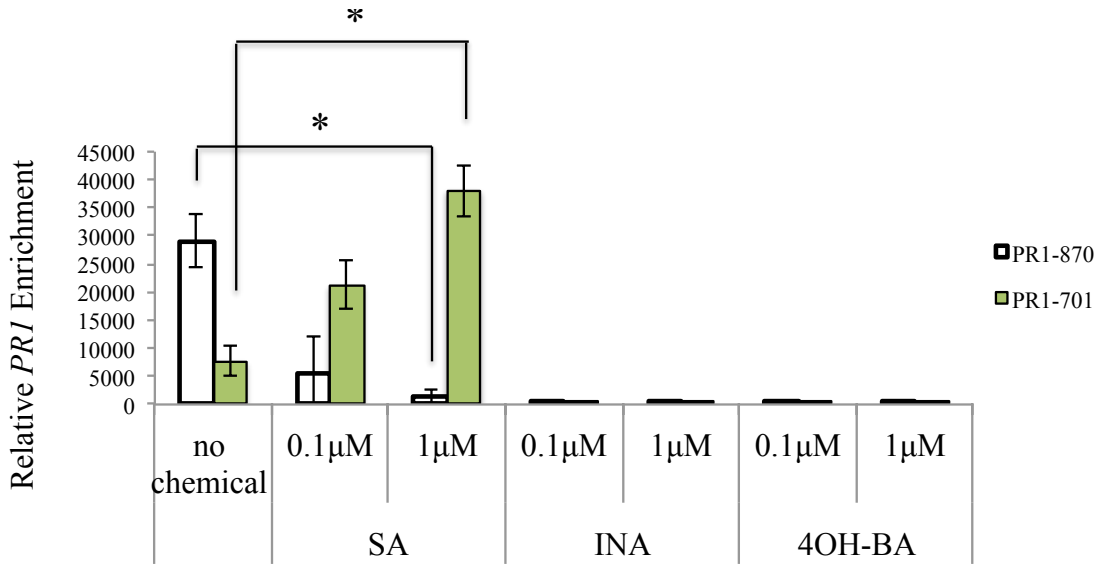


Figure 3. Positive SA regulation of NPR1 binding to the PR1 promoter and negative INA and 4-OH-BA regulations of NPR1 binding to the PR1 promoter

A. Schematic representation of PR1 promoter deletion constructs used for modified in vitro ChIP experiments. The arrow indicates the start of transcription. Each box represents a fragment of the PR1 promoter sequence cloned into a plasmid and used as DNA source in modified ChIP experiments. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). White box indicates the -1 to -870bp fragment of the PR1 promoter, which sequences from -870bp to -1023bp deleted (Grey colour box). Green box represents the -1 to -701bp fragment of the PR1 promoter that sequences from -701bp to -1023bp deleted (Grey colour box). Blue box illustrates the -1 to -621bp fragment of the PR1 promoter, which sequences from -621bp to -1023bp deleted (Grey colour box). The last white box shows full-length NPR1 protein (amino acids 1-593).

B. Structures of salicylic acid (SA) and its inactive analogues, 2,6 Dichloro INA and 4-Hydroxy Benzoic acid (4-OH-BA).

C. Histogram showing modified ChIP experiments performed on the -870, -701, and -620 fragments of the PR1 promoter with NPR1 protein after addition of 10 and 80µM of SA, confirming that SA positively regulates binding of NPR1 to the PR1 promoter. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

D. Histogram showing modified ChIP experiments on the -870 and -701 fragments of the PR1 promoter with NPR1 protein after addition of 1 and 0.1µM INA and 4-OH-BA, confirming that they negatively regulate binding of NPR1 to the PR1 promoter. SA was used as a positive control measure. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

4.4. Positive regulation of NPR1 binding by SA to the -636 to -646 fragment of the PR1 promoter

Upon confirmation that salicylic acid (SA) positively affects the recruitment of the NPR1 protein to the -701 fragment of the PR1 promoter (Figure 3C and 3D, green bars), research was directed towards further specifying the binding site location of NPR1 within the -701 fragment (-1 to -701bp). To accomplish this, modified ChIP experiments were conducted to test NPR1 recruitment to different PR1 promoter deletion constructs made within the -701fragment region (Figure 4A).

The obtained results revealed that in the presence of SA (1μM), the -621 to -674 base pair fragment of the PR1 promoter contained significant regions for NPR1 recruitment. A student two-tailed t-test showed that the amount of relative *PR1* enrichment in the -674 fragment was similar to values obtained for the -701 fragment (p-value is 0.2600; p-value < 0.05), with both fragments yielding a higher degree of relative *PR1* enrichment after addition of SA. However, the amount of relative *PR1* enrichment in the -621 fragment was found to be significantly smaller than the amounts obtained for the -701 fragment (p-value is 0.0142; p-value < 0.05), both in the presence and absence of SA.

Based on these results, a new experiment was conducted to further specify the binding site of NPR1 within the -621 to -674 regions. As a first step, new deletion constructs had to be chosen for a series of new ChIPs experiments. Past work conducted by Lebel et al. (1998) on PR1 promoter deletion analysis showed that two 10 base pair linker-scanning mutations at the -610 to -640 bp location upstream of the transcription initiation site of the PR1 promoter are very important for *PR1* gene expression in *Arabidopsis*. Particularly, mutation in LS5 induced PR1 expression but mutation in LS7 abolished PR1 expression. More importantly, TGA2 has been shown

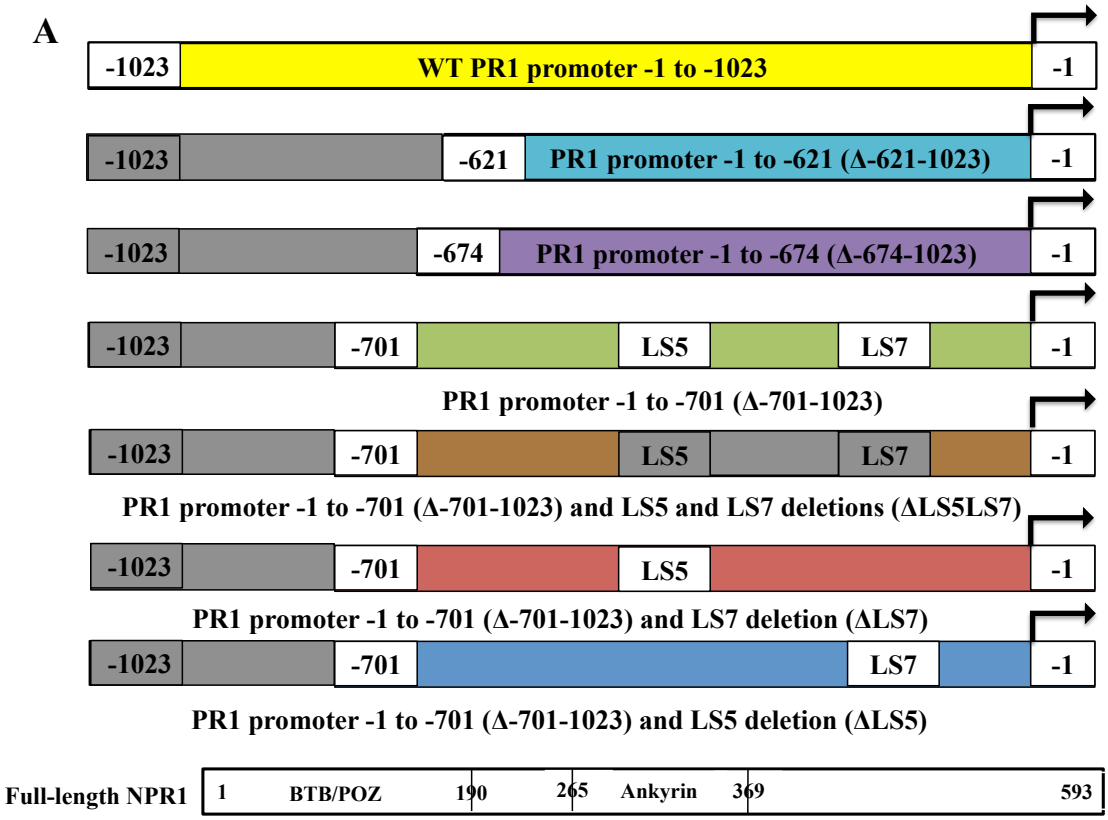
to also bind to LS5 (-666 to -676 bp) and LS7 (-636 to -646 bp) (Boyle et al., 2009). Accordingly, another series of PR1 deletion constructs were prepared around the -635 to -675 base pair sequence within the -701 fragment of the PR1 promoter, including LS5LS7 deletions (base pair -635 to -675 is deleted), only LS5 deletion (base pair -666 to -676 is deleted), and only LS7 deletion (base pair -636 to -646 is deleted).

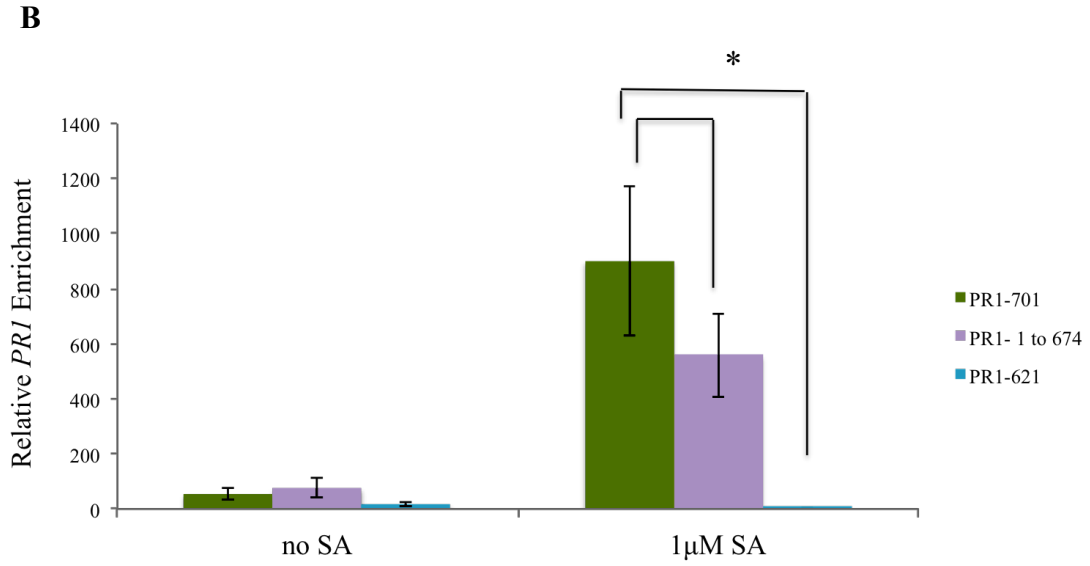
Next, we tested NPR1 binding to the PR1 promoter LS5LS7 deletions (Figure 4A, brown box, LS7 and LS5 deletions) in the presence of 1μM SA. Findings indicated that the mutation was sufficient enough to abolish relative *PR1* enrichment in the presence and absence of SA (Figure 4C, brown bar) with a significant difference in relative *PR1* enrichment observed in comparison to the -701 fragment, where both LS5 and LS7 were present (Student two-tailed t-test, p-value is 0.0016; p-value < 0.05) (Figure 4C, green bars).

Next, NPR1 binding to the PR1 promoter only LS7 and only LS5 deletions were tested (Figure 4A, red and blue box respectively). Surprisingly, results obtained for the LS7 deletion showed a complete loss of relative *PR1* enrichment in the presence of SA, with values similar to those obtained for LS5LS7 deletions (Figure 4A, red bar and Figure S3A). A student two-tailed t-test revealed that a significant decrease in relative *PR1* enrichment occurred in the presence of SA for the LS7 deletion compared to the -701 fragment, which contains both LS5 and LS7 (Figure 4C, green bar) (p-value is 0.00208; p-value < 0.05). Moreover, the obtained relative *PR1* enrichment amount in the presence of SA for the deletion construct containing the only LS5 deletion (Figure 4C, blue bar) was found to be similar to values obtained for the relative *PR1* enrichment of the -701 fragment, which contains both LS5 and LS7 (Figure 4C, green bar) A student t-test confirmed that there were no significant differences in enrichment values between the LS5 deletion construct and the -701

fragment (p-value is 0.0966; p-value < 0.05), and that both showed a similar increase in enrichment after addition of SA.

Briefly, these results suggest that in the presence of SA, NPR1 is recruited to the PR1 promoter around the -636 to -646 base pair sequence (LS7 site).


A
-1023
WT PR1 promoter -1 to -1023
-1
-621
PR1 promoter -1 to -621 (Δ-621-1023)
-674
PR1 promoter -1 to -674 (Δ-674-1023)
-701
LS5
LS7
PR1 promoter -1 to -701 (Δ-701-1023)
PR1 promoter -1 to -701 (Δ-701-1023) and LS5 and LS7 deletions (ΔLS5LS7)
PR1 promoter -1 to -701 (Δ-701-1023) and LS7 deletion (ΔLS7)
PR1 promoter -1 to -701 (Δ-701-1023) and LS5 deletion (ΔLS5)
Full-length NPR1
1
BTB/POZ
190
265
Ankyrin
369
593


B
*
1400
1200
1000
800
600
400
200
0
Relative PR1 Enrichment
no SA
1μM SA
PR1-701
PR1- 1 to 674
PR1-621

C

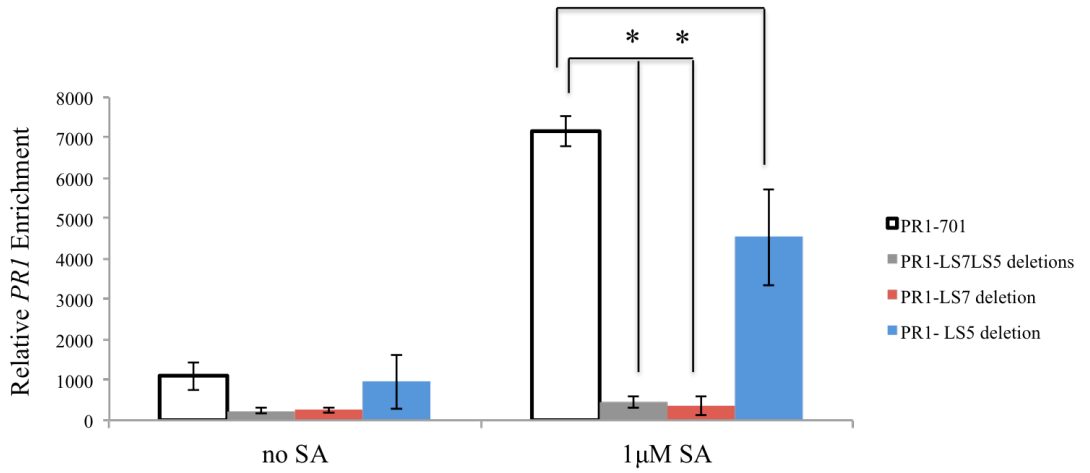


Figure 4. Binding of NPR1 to the PR1 promoter at the -636 to -646 base pair sequence in the presence of SA

A. Schematic representation of PR1 promoter deletion constructs used for modified in vitro ChIP experiments. The arrow indicates the start of transcription. Each box represents a fragment of the PR1 promoter sequence cloned into a plasmid and used as DNA source in the modified ChIP experiment. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). The light blue box indicates the -1 to -621bp fragment of the PR1 promoter sequence that sequences from -621bp to -1023bp deleted (Grey colour box). The purple box indicates the -1 to -674bp fragment of the PR1 promoter sequence, which sequences from -674bp to -1023bp deleted (Grey colour box). Green box represents the -1 to -701bp fragment of the PR1 promoter that sequences from -701bp to -1023bp deleted (Grey colour box). The brown box shows the -701 fragment of the PR1 promoter sequence, where both LS5 (-666 to -676 bp) and LS7 (-636 to -646 bp) sequences deleted (Grey colour box). The red box indicates the -701 fragment of the PR1 promoter sequence, where only the LS7 (-636 to -646 bp) sequence deleted. The blue box indicates the -701 fragment of the PR1 promoter sequence, where only the LS5 (-666 to -676 bp) sequence deleted. The white box indicates full-length NPR1 protein (amino acids 1-593).

B. Histogram illustrating modified ChIP experiments performed on PR1 promoter deletion constructs, confirming that in presence of SA, the -621 to -674bp sequence region of the PR1 promoter yielded the most *PR1* enrichment. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment amounts were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

C. Histogram illustrating modified ChIP experiments performed on PR1 promoter deletion constructs, confirming that NPR1 binds to the LS7 (-636 to -646 bp) fragment of the PR1 promoter in presence of SA. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05)**.**

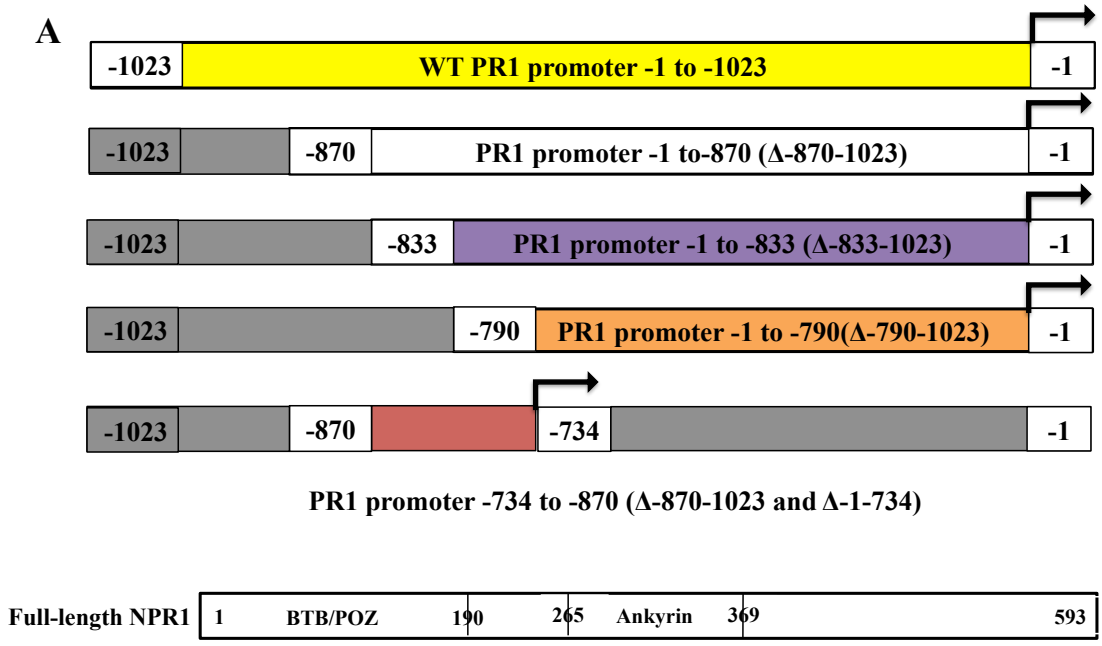
4.5. Binding of NPR1 to the -790 to -833 fragment region of the PR1 promoter in the absence of SA

Similar to the analyses conducted to further specify the binding site within the -701 fragment of the PR1 promoter, experiments were conducted to further specify the NPR1 binding site within the -870 fragment of the PR1 promoter. The obtained results (Figure 3C and Figure 3D) demonstrated a low relative *PR1* enrichment factor in the presence of SA, confirming that NPR1 is recruited to the -870 fragment of the PR1 promoter in the absence of SA. Next, modified ChIP experiments were conducted to test NPR1 binding to a series of PR1 deletion constructs, including -1 to -870bp, -1 to -790bp, -1 to -833bp, and -734 to -870bp fragments (Figure 5A and Figure S2) in the presence and absence of 1μM SA.

The obtained findings established that the obtained amounts of relative *PR1* enrichment were not statistically different between the -870 and -833 fragments of the PR1 promoter in the absence of SA (Student two-tailed t-test, p-value is 0.6501; p-value < 0.05) (Figure 5B, white and purple bars). In addition, both fragments showed a higher degree of relative *PR1* enrichment in the absence of SA, followed by a complete loss of enrichment after addition of SA. Remarkably, the relative *PR1* enrichment amounts for the -1 to -870bp and -734bp to -870 fragments were found to be statistically similar (Student two-tailed t-test, p-value is 0.1183; p-value < 0.05) (Figure 5B, white and red bars) and also for the -1 to -833bp and -734 to -870 fragments (Student two-tailed t-test, p-value is 0.1158; p-value < 0.05) (Figure 5B, purple and red bars), with a higher degree of enrichment observed prior to addition of SA. These results suggest that the -734bp to -833 fragment region of the PR1 promoter could potentially contain an important site for NPR1 binding. However, upon inspection of the -790 fragment results, it was observed that relative *PR1*

enrichment was completely abolished (Figure 5B, orange bars) (Figure S3B), and that a significant difference existed in comparison to relative *PR1* enrichment values obtained for the -870 fragment (Student two-tailed t-test, p-value is 0.0148; p-value < 0.05).

In conclusion, our results confirmed that in the absence of SA, NPR1 recruitment occurs between the -790 to -833 fragments of the PR1 promoter.



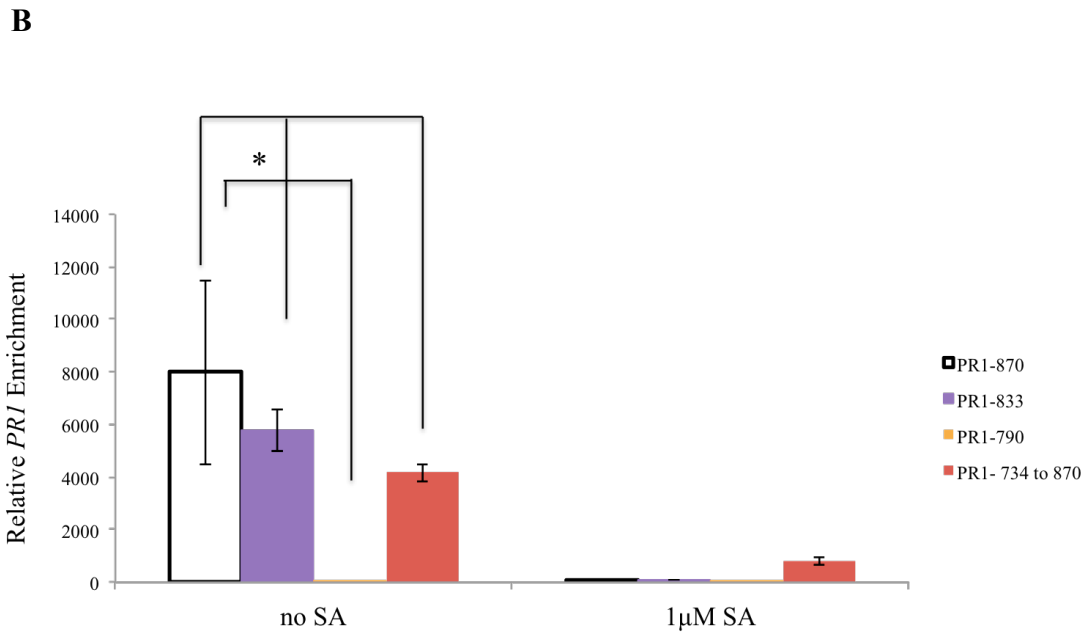


Figure 5. Binding of NPR1 to the -790 to -833 fragment of the PR1 promoter in the absence of SA

A. Schematic representation of PR1 promoter deletion constructs used for modified in vitro ChIP experiment. The arrow indicates the start of transcription. Each box represents a fragment of the PR1 promoter sequence cloned into a plasmid and used as DNA source in the modified ChIP experiment. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). White box indicates the -1 to -870bp fragment of the PR1 promoter, which sequences from -870bp to -1023bp

deleted (Grey colour box). Purple box indicates the -1 to -833bp fragment of the PR1 promoter, which sequences from -833bp to -1023bp deleted (Grey colour box). Orange box indicates the -1 to -790bp fragment of the PR1 promoter, which sequences from -790bp to -1023bp deleted (Grey colour box). Red box indicates the -734 to -870bp fragment of the PR1 promoter, which sequences from -1bp to -734bp and -870bp to -1023bp deleted (Grey colour box). The last white box shows full-length NPR1 protein (amino acids 1-593).

B. Histogram illustrating modified ChIP experiments performed on PR1 promoter deletion constructs, confirming that NPR1 binds to the -790 to -833 fragment of the PR1 promoter in the absence of SA. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

4.6 Determination of NPR1 DNA binding domains in both the N and C-terminus

After successful determination of NPR1 binding sites on the PR1 promoter (Figure 4 and 5), the location of the DNA binding domain was subjected to investigation. Considering our past findings, which determined that NPR1 was recruited at two different sites on the PR1 promoter in both the presence and absence of SA (Figure 4 and 5), it was hypothesized that NPR1 might also have two distinct DNA binding domains that facilitate its recruitment to the PR1 promoter. The previously discussed findings regarding NPR1's distinct binding preference in the presence and absence of SA (Figure 3, 4, and 5) further validated the initial hypothesis.

In related work conducted by Rochon et al. (2006), the amino acids cysteine 521 and 529 ($Cys^{521/529}$) of the NPR1 protein were determined to be necessary for *PR1* gene activation in vivo, as well as for the transactivation function of the full-length NPR1 and for NPR1 Δ1-513 (NPR1 513-593). Further, Wu et al. (2012) demonstrated that SA binds NPR1 specifically through $Cys^{521/529}$ and a Copper transition metal. Together, these findings confirm the importance of the NPR1 C-terminus region for SA binding to NPR1, and consequently, to the PR1 promoter; accordingly, an analysis of the C-terminus region was conducted to identify a potential DNA binding domain.

4.7 Location of DNA binding domain of NPR1 for binding to the -701 fragment of the PR1 promoter in the presence of SA

Tests involving the binding of a series of NPR1 N-terminus deletion constructs to the -701 fragment of the PR1 promoter, both in the presence and absence of SA, were conducted to identify the potential DNA binding domain that recruits NPR1 to the PR1 promoter in presence of SA. All deletion constructs were made to NPR1 Δ1-374, meaning that they lack the amino acids 1-374, with the exception of NPR1 Δ1-513 (NPR1 513-593) (Figure 6A). Full-length NPR1 was recruited as a positive control measure. Amino acid 374 is located right after ankyrin repeats of NPR1 as predicted by Pfam (Finn et al., 2006) and SMART (Letunic et al., 2006). Amino acid 513 on C-terminus region corresponds to the beginning of the last stretch of negatively charged and hydrophobic residues (Cress and Triezenberg, 1991).

Consistent with our previous results (Figure 3 and 4), relative *PR1* enrichment amounts for full-length NPR1 were found to be statistically greater in the presence of SA (Student two-tailed t-test, p-value is 0.0402; p-value < 0.05) (Figure 6B). Similar to the results obtained for full-length NPR1, relative *PR1* enrichment amounts for NPR1 374-593 (NPR1 Δ1-374) were also found to be statistically greater after addition of SA (Student two-tailed t-test, p-value is 0.0279; p-value < 0.05) (Figure 6B). On the other hand, no significant differences were found in relative *PR1* enrichment amounts obtained before and after addition of SA for both NPR1 374-535 (NPR1 Δ1-374 and Δ535-593) (Student two-tailed t-test, p-value is 0.4855; p-value < 0.05) and NPR1 513-593 (NPR1 Δ1-513) (Student two-tailed t-test, p-value is 0.3668; p-value < 0.05) (Figure 6B). Together, these results suggest that a potential DNA binding domain could be localized in the C-terminus region between amino acids 374-593.

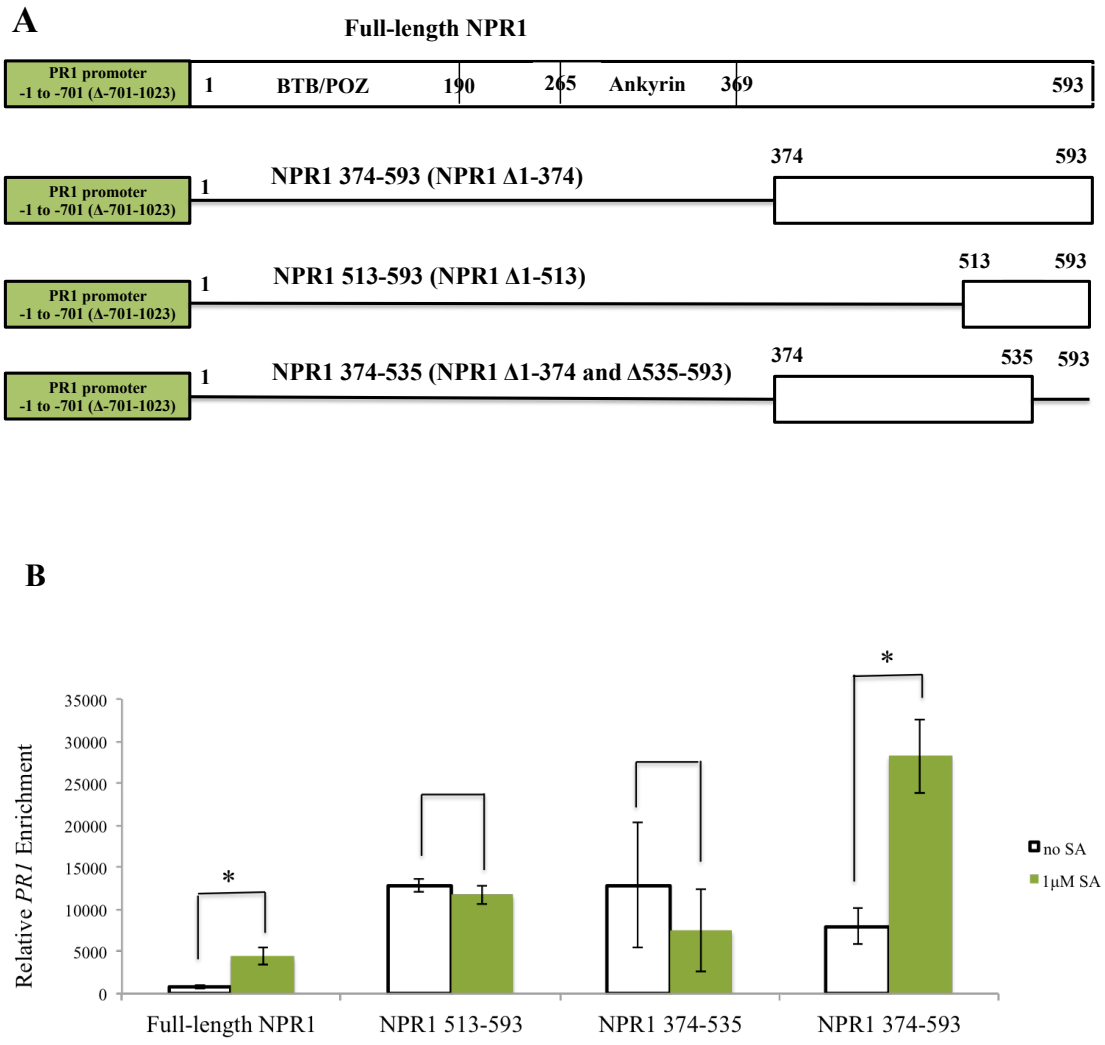


Figure 6. Possible location of the DNA binding domain of NPR1 for binding to the -701 fragment of the PR1 promoter in the presence of SA

A. Schematic representation of NPR1 protein deletion constructs. All constructs were made to NPR1 Δ1-374, meaning that they lack the amino acids 1 to 374, with the exception of NPR1 Δ1-513 (NPR1 513-593). The first box from the top shows full-length NPR1 (amino acid 1-593). The second box indicates NPR1 Δ1-374 that amino acid 1 to 374 deleted. The third box represents NPR1 Δ1-513 that amino acid 1 to 513 deleted. The last box illustrates NPR 374-535 (NPR1 Δ1-374 and Δ535-593), which amino acids 1 to 374 and 535 to 593 deleted. Green box shows the -1 to -701 fragment of the PR1 promoter that is used as DNA source, along with NPR1 deletion constructs in ChIP experiments.

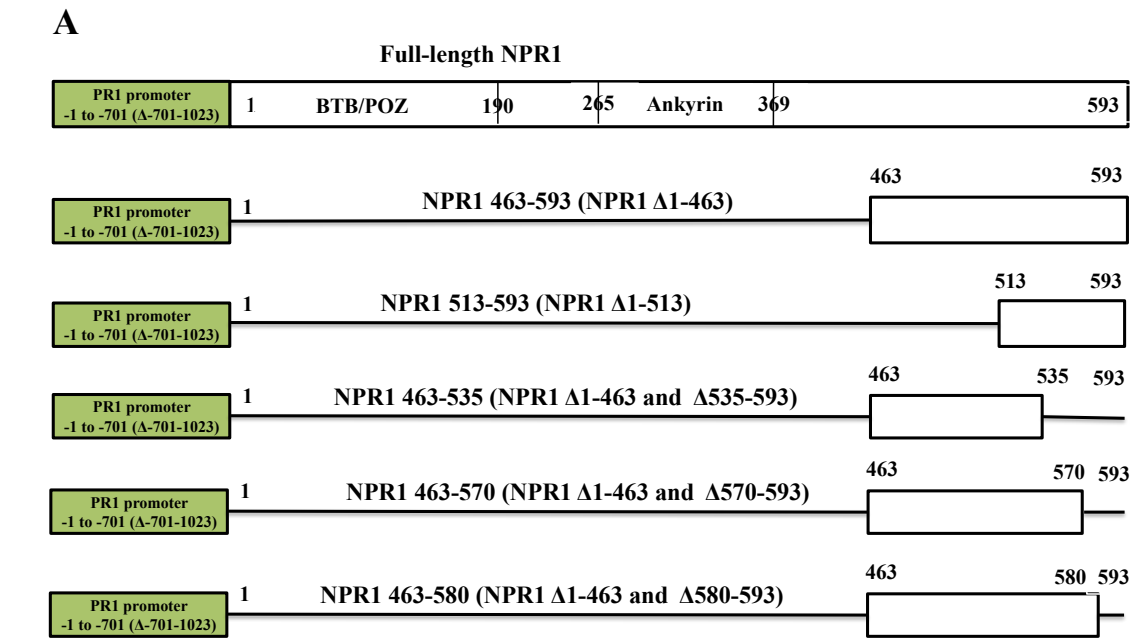
B. Histogram illustrating the results of modified ChIP experiments performed on NPR1 deletion constructs and the -701 fragment of the PR1 promoter after addition of SA, suggesting possible localization of the NPR1 DNA binding domain between amino acids 374 to 593. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

In order to further narrow down the localization of the DNA binding domain for the binding of NPR1 to the -701 fragment of the PR1 promoter, which was previously suggested through our findings to be within the amino acids 374-593 (Figure 6), a new series of experiments involving new deletion constructs were conducted. For this purpose, a new series of NPR1 N-terminus deletion constructs, made from NPR1 Δ1-463 (NPR1 463-593, i.e., lacking amino acids 1 to 463) (Figure 7A) were tested for binding to the -701 fragment of the PR1 promoter in the presence of SA. Full-length NPR1 and NPR1 513-593 (Δ1-513) were recruited as positive control measures. Amino acid 463 is located on C-terminus region after ankyrin repeats and is the end point of sequence similarity with *Drosophila* ankyrin 2 (GenBank accession number AAN12046.1) (Rochon et al., 2006).

Consistent with our previous results (Figure 3, 4, and 6B), relative *PR1* enrichment values were observed to be statistically greater for full-length NPR1 in the presence of SA (Student two-tailed t-test, p-value is 0.0505; p-value < 0.05) (Figure 7B). No statistically significant differences were found for relative *PR1* enrichment values obtained before and after SA addition for both NPR1 463-535 (NPR1 Δ1-463 and Δ535-593) (Student two-tailed t-test, p-value is 0.4855; p-value < 0.05) and NPR1 513-593 (NPR1 Δ1-513) (Student two-tailed t-test, p-value is 0.1328; p-value < 0.05) (Figure 7B). In addition, there were no statistically significant differences found for relative *PR1* enrichment values before and after addition of SA for both NPR1 463-593 (NPR1 Δ1-463) (Student two-tailed t-test, p-value is 0.0580; p-value < 0.05) and NPR1 463-570 (NPR1 Δ1-463 and Δ570-593) (Student two-tailed t-test, p-value is 0.0277; p-value < 0.05) (Figure 7B), with both deletion constructs yielding a higher degree of relative *PR1* enrichment after addition of SA. On the other hand, for NPR1 463-580 (NPR1 Δ1-463 and Δ580-593), a statistically significant difference was

found in relative *PR1* enrichment before and after addition of SA (Student two-tailed t-test, p-value is 0.0418; p-value < 0.05), with an observed reduction in relative *PR1* enrichment after addition of SA (Figure 7B).

In conjunction, the obtained finding suggest that in the presence of SA, the potential DNA binding domain of NPR1 DNA for binding to the -701 fragment of the PR1 promoter is localized between amino acids 463-570.



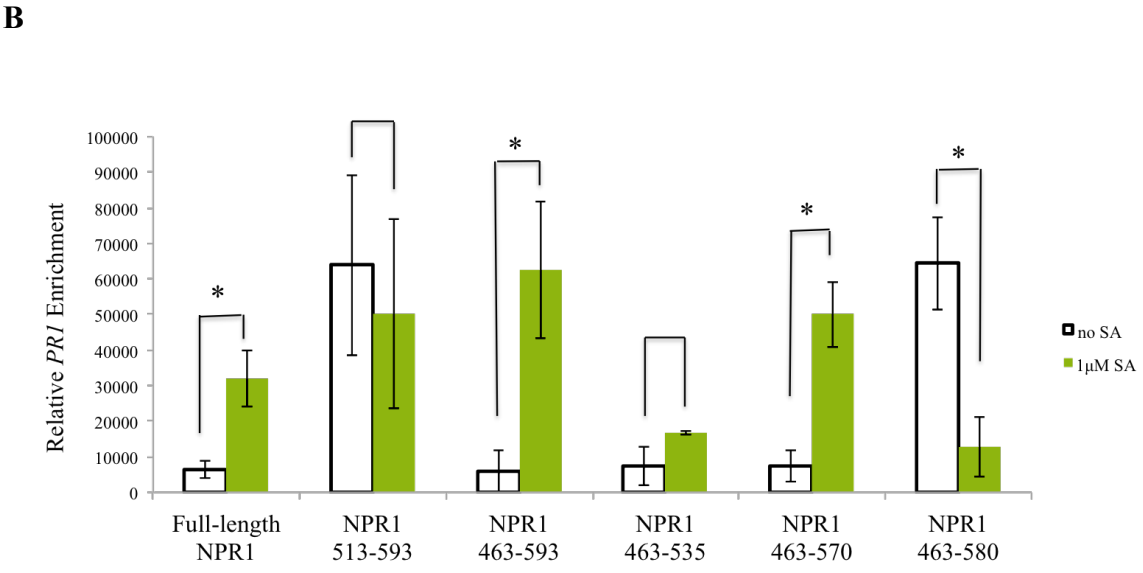


Figure 7. Possible location of the DNA binding domain of NPR1 for binding to -701 fragment of the PR1 promoter in the presence of SA

A. Schematic representation of NPR1 deletion constructs. All constructs were made to NPR1 Δ1-463, meaning that they lack amino acids 1 to 463, with the exception of NPR1 Δ1-513 (NPR1 513-593). The first box from the top shows full-length NPR1 (amino acid 1-593). The second box indicates NPR1 Δ1-463, which amino acid 1 to 463 deleted. The third box represents NPR1 Δ1-513 that amino acid 1 to 513 deleted. The fourth box illustrates NPR 463-535 (NPR1 Δ1-463 and Δ535-593), which amino acids 1 to 463 and 535 to 593 deleted. The fifth box shows NPR1 463-570 (NPR1 Δ1-463 and Δ570-593) that amino acids 1 to 463 and 570 to 593 deleted. The last box indicates NPR1 463-580 (NPR1 Δ1-463 and Δ580-593), which amino acids 1 to 463 and 580 to 593 deleted. Green box shows

the -1 to -701 fragment of the PR1 promoter that is used as DNA source along with NPR1 deletion constructs used in ChIP experiments.

B. Histogram illustrating modified ChIP experiments performed on NPR1 C-terminus deletion constructs and the -701 fragment of the PR1 promoter after addition of SA, suggesting possible location of the NPR1 DNA binding domain to be between amino acids 463 to 570. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

Upon determination of the approximate location for the DNA binding domain site for NPR1 to be between amino acids 463-570 (Figure 7), further investigations were needed to further specify the site location, as the region still spanned 107 amino acids in length. As explained earlier, all used NPR1 N-terminus deletion constructs were made either to amino acid 374 (Figure 6A) or amino acid 463 (Figure 7A), meaning that they either lacked amino acids 1-374 or 1-463, respectively. As such, despite findings that pointed towards a significant difference in some of the tested NPR1 N-terminus deletion constructs, including NPR1 374-593 (Figure 6) and NPR1 463-570 (Figure 7), a conclusive determination could not be attained. This peculiarity could be due to improper protein folding, which can be attributed to the short length of the proteins and disruption of balances of forces that contribute to protein folding and stabilization (Fersht 1993; Lazaridis et al. 1995; Pace et al. 1996). Difficulties in accurately determining the DNA binding site from these constructs prompted the preparation of new NPR1 N-terminus deletion constructs made to the full-length NPR1. Application of this approach eliminated the possibility of improper protein folding hindering a definite conclusion. For these experiments, the binding of all newly-made NPR1 N-terminus deletion constructs to the -701 fragment of the PR1 promoter were tested in the presence of SA (Figure 8A).

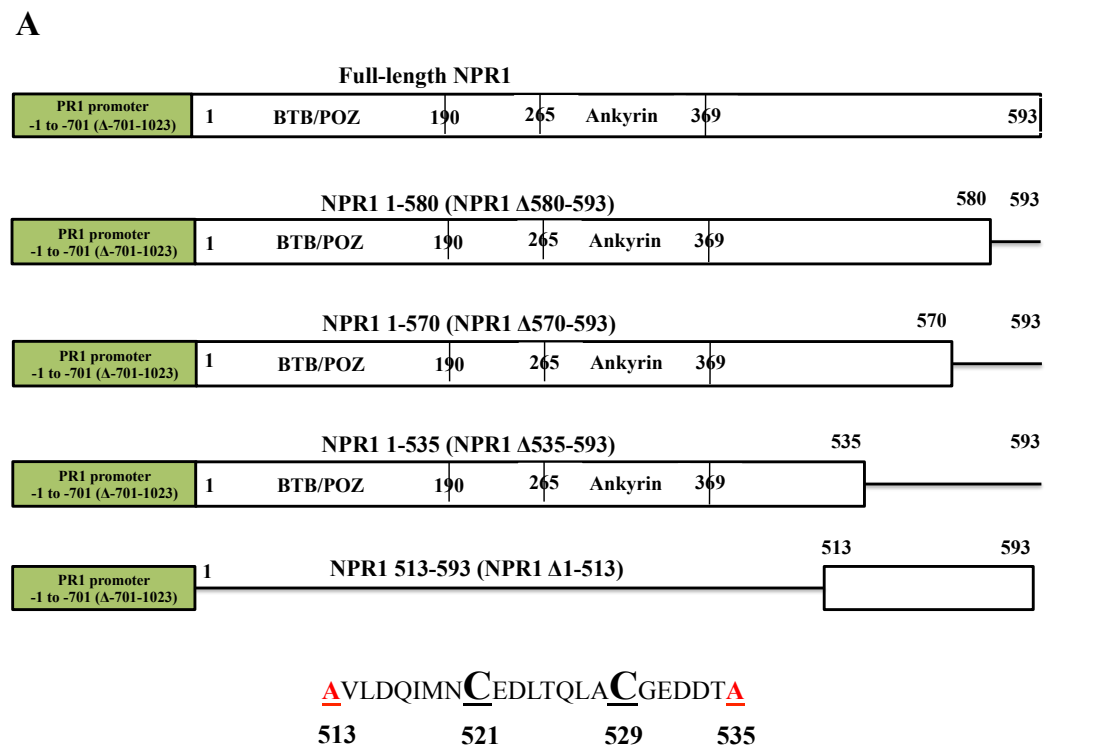
A statistically significant difference was found to exist between relative *PR1* enrichment values for both NPR1 1-570 (NPR1 Δ570-593) and NPR1 1-580 (NPR1 Δ580-593) deletion constructs before and after addition of SA (Student two-tailed t-test, p-values are 0.0409 and 0.0335 respectively; p-value < 0.05). A higher degree of relative *PR1* enrichment was observed for NPR1 1-570 (NPR1 Δ570-593) after addition of SA, while a lesser degree of relative *PR1* enrichment was observed for NPR1 1-580 (NPR1 Δ580-593) after addition of SA (Figure 8B). These results were

consistent with our previous findings, as a higher degree of *PR1* enrichment was also found for NPR1 463-570 (NPR1 Δ1-463 and Δ570-593) after addition of SA, while NPR1 463-580 (NPR1 Δ1-463 and Δ580-593) showed a lesser degree of *PR1* enrichment after addition of SA (Figure 8B).

Consistent with our previous results (Figure 6B, 7B), Full-length NPR1 yielded a more statistically significant increase in relative *PR1* enrichment after addition of SA (Student two-tailed t-test, p-value is 0.0206; p-value < 0.05) (Figure 8B). Relative *PR1* enrichment for NPR1 513-593 (NPR1 Δ1-513) was also consistent with previously obtained results (Figure 6B, 7B), with no statistically significant differences found before and after addition of SA (Student two-tailed t-test, p-value is 0.6636; p-value < 0.05) (Figure 8B). Surprisingly, a statistically significant increase in relative *PR1* enrichment for the NPR1 1-535 (NPR1 Δ535-593) construct was found after addition of SA (Student two-tailed t-test, p-value is 0.0406; p-value < 0.05) (Figure 8B). Briefly, the obtained results suggest that in the presence of SA, the DNA biding domain of NPR1 for binding to the -701 fragment of the PR1 promoter is localized between amino acids 513-535.

In conclusion, our findings, as can be seen from Figures 6, 7, and 8, demonstrated that the relative *PR1* enrichment for the NPR1 513-593 (NPR1 Δ1-513) C-terminus deletion construct was consistent, with no observed differences in enrichment values before and after addition of SA. Similar consistency was also observed for full-length NPR1, where a statistically larger amount of relative *PR1* enrichment was observed after addition of SA. Successful and accurate determination of the DNA binding domain was finalized by the results obtained for the NPR1 1-535 (NPR1 Δ535-593) C-terminus deletion construct (Figure 8B), for which a significantly higher relative *PR1* enrichment amount was observed after addition of

SA. A comparison of relative *PR1* enrichment values between NPR1 513-593 (NPR1 Δ1-513) and NPR1 1-535 (NPR1 Δ535-593) deletion constructs (Figure 8B) finalized that in the presence of SA, the DNA binding domain site for NPR1 binding to the -701 fragment of the PR1 promoter lies between amino acids 513-535.



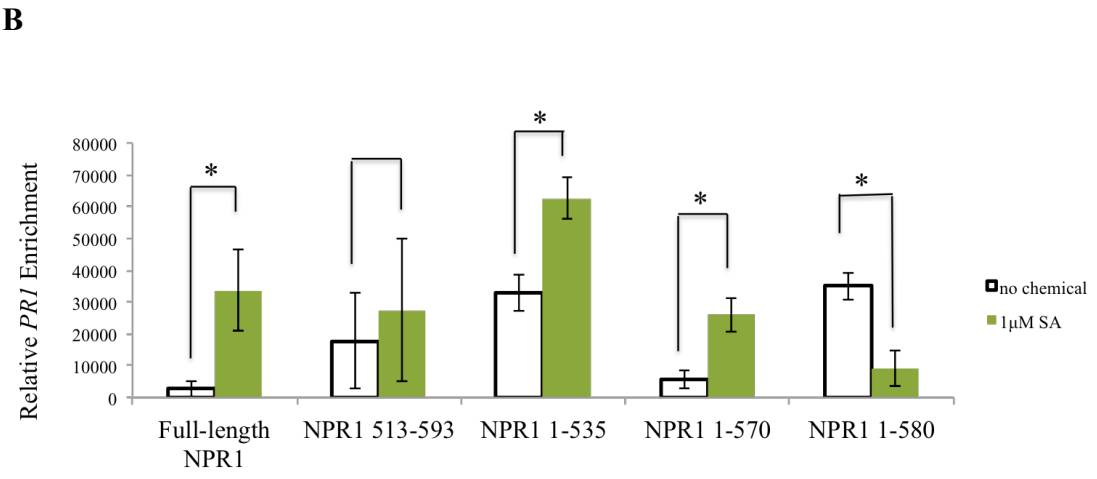


Figure 8. Possible location of the DNA binding domain of NPR1 for binding to the -701 fragment of the PR1 promoter in the presence of SA

A. Schematic representation of NPR1 deletion constructs. All constructs were made to full-length NPR1, meaning that they all start at amino acid 1, with the exception of NPR1 Δ1-513 (NPR1 513-593). The first box from the top illustrates the full-length NPR1 (amino acid 1-593). The second box indicates NPR11-580 (NPR1 Δ580-593) that amino acid 580 to 593 deleted. The third box represents NPR11-570 (NPR1 Δ570-593), which amino acid 570 to 593 deleted. The fourth box illustrates NPR 1-535 (NPR1 Δ535-593) that amino acid 535 to 593 deleted. The fifth box shows NPR1 513-593 (NPR1 Δ1-513), which amino acid 1 to 513 deleted. The green box shows the -1 to -701 fragment of the PR1 promoter used as a DNA source, along with NPR1 deletion constructs used in ChIP

experiments. Sequences of NPR1 C-terminus region (amino acids 513-535) show Cys^{521} and Cys^{529} positions.

B. Histogram illustrating modified ChIP experiments performed on NPR1 deletion constructs and the -701 fragment of the PR1 promoter after SA addition, confirming localization of the NPR1 DNA binding domain between amino acids 513 to 535. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

4.8 Location of DNA binding domain of NPR1 for binding to the -870 fragment of PR1 promoter in the absence of SA

After successful determination of the DNA binding domain site of NPR1 that allows its recruitment to the -701 fragment of the PR1 promoter in the presence of SA (Figure 6,7, and 8), the DNA binding domain site that facilitates NPR1 recruitment to the -870 fragment of PR1 promoter remained to be elucidated. Similar to our previous approach, the recruitment of a series of NPR1 N-terminus deletion constructs to the -870 fragment of the PR1 promoter were tested only in the absence of SA, using modified ChIPs. All deletion constructs were made to NPR1 Δ1-374, meaning that they lacked amino acids 1-374, with the exception of NPR1 513-593 (NPR1 Δ1-513) and NPR1 533-593 (NPR1 Δ1-533) (Figure 9A). Full-length NPR1 was recruited as a positive control measure. Amino acid 374 is located right after ankyrin repeats of NPR1 as predicted by Pfam (Finn et al., 2006) and SMART (Letunic et al., 2006). Amino acid 513 on C-terminus region corresponds to the beginning of the last stretch of negatively charged and hydrophobic residues, a signature of transactivation domains (Cress and Triezenberg, 1991). Amino acid 533 is localized on C-terminus region right before the nuclear localization signal (Kinkema et al., 2000).

As expected, *PR1* enrichment was not observed for all tested NPR1 N-terminus deletion constructs in the absence of SA. Full-length NPR1, conversely, showed a significantly higher degree of relative *PR1* enrichment compared with other C-terminus deletion constructs, including NPR1 374-580 (NPR1 Δ1-374 and NPR1 Δ580-593), NPR1 374-570 (NPR1 Δ1-374 and NPR1 Δ570-593), NPR1 374-535 (NPR1 Δ1-374 and NPR1 Δ535-593), NPR1 374-593 (NPR1 Δ1-374), NPR1 533-593 (NPR1 Δ1-533), and NPR1 513-593 (NPR1 Δ1-513) (Figure 9B) (Student two-tailed t-test,

p-values are 0.0082, 0.0048, 0.0101, 0.0079, 0.0107, and 0.0039, respectively; p-value < 0.05).

Together, these results demonstrated that the C-terminus region of NPR1 is not associated with its recruitment to the -870 fragment of the PR1 promoter in the absence of SA, further confirming our previous results (Figure 6,7, and 8) regarding the importance of the C-terminus region for binding to SA and the -701 fragment of the PR1 promoter.

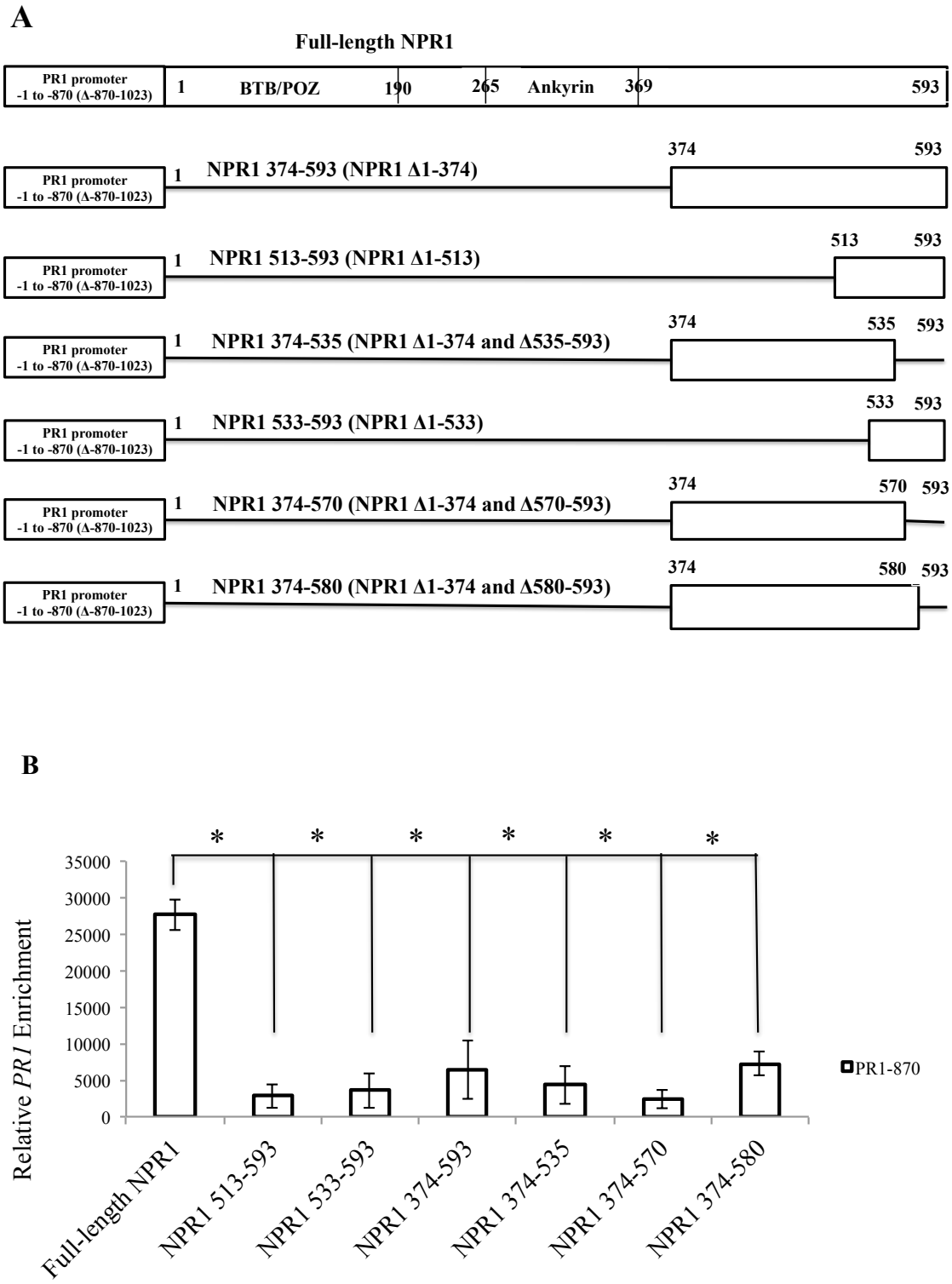


Figure 9. Irrelevance of C-terminus region of NPR1 for binding to the -870 fragment of the PR1 promoter in absence of SA

A. Schematic representation of NPR1 deletion constructs. All constructs were made to NPR1 Δ1-374, meaning that they lack the amino acid 1 to 374 except for NPR1 Δ1-513 (NPR1 513-593) and NPR1 Δ1-533 (NPR1 533-593). The first box from the top illustrates the

full-length NPR1 protein (amino acids 1-593). The second box indicates NPR1 374-593 (NPR1 Δ1-374) that amino acid 1 to 374 deleted. The third box represents NPR1 513-593 (NPR1 Δ1-513), which amino acid 1 to 593 deleted. The fourth box represents NPR1 374-535 (NPR1 Δ1-374 and NPR1 Δ535-593), which amino acids 1 to 374 and 535 to 593 deleted. The fifth box shows NPR1 533-593 (NPR1 Δ1-533) that amino acids 1 to 533 deleted. The sixth box illustrates NPR 374-570 (NPR1 Δ1-374 and NPR1 Δ570-593), which amino acids 1 to 374 and 570 to 593 deleted. Last box shows NPR 374-580 (NPR1 Δ1-374 and NPR1 Δ580-593) that amino acids 1 to 374 and 580 to 593 deleted. The white box shows the -1 to -870 fragment of the PR1 promoter that is used as DNA in ChIP experiments.

B. Histogram illustrating modified ChIP experiments performed on NPR1 C-terminus deletion constructs and the -870 fragment in the absence of SA, confirming that the C-terminus is not involved in NPR1 binding to the -870 fragment. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment amounts were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

In accordance with our findings, the C-terminus region of NPR1 was concluded to not be involved in the recruitment of NPR1 to the -870 fragment of the PR1 promoter (Figure 9). Consequently, binding of the NPR1-POZ domain (NPR1 1-190 (Δ190-593), located in N-terminus region) to the -870 fragment of the PR1 promoter was tested in the presence and absence of SA (Figure 10A). Full-length NPR1 was recruited as a positive control, while two N-terminus region deletion constructs, NPR1 374-593 (NPR1 Δ1-374) and NPR1-Δ513 (NPR1 Δ1-513), were used as negative controls.

Interestingly, a similar amount of relative *PR1* enrichment was found for NPR1 1-190 (Δ190-593) and full-length NPR1 in the absence and presence of SA (Figure 10B) (Student two-tailed t-test, p-values are 0.2390 and 0.5394respectively; p-value < 0.05), with a higher degree of *PR1* enrichment occurring in the absence of SA. As expected, both C-terminus region deletion constructs, NPR1 374-593 (Δ1-374) and NPR1 513-593 (Δ1-513), displayed significantly lesser amounts of relative *PR1* enrichment both in the absence and presence of SA when compared to full-length NPR1 (Figure 10B). A Student two-tailed t-test assessment on relative *PR1*enrichment values obtained in the absence of SA for NPR1 513-593 (Δ1-513) and NPR1 374-593 (Δ1-374) presented p-values of 0.0085 and 0.0084, respectively.

Together, these findings further confirmed that the C-terminus region of NPR1 is not involved in its recruitment to the -870 fragment of the PR1 promoter in the absence of SA, and showed that a possible DNA binding domain may be located in the N-terminus region, between amino acids 1-190.

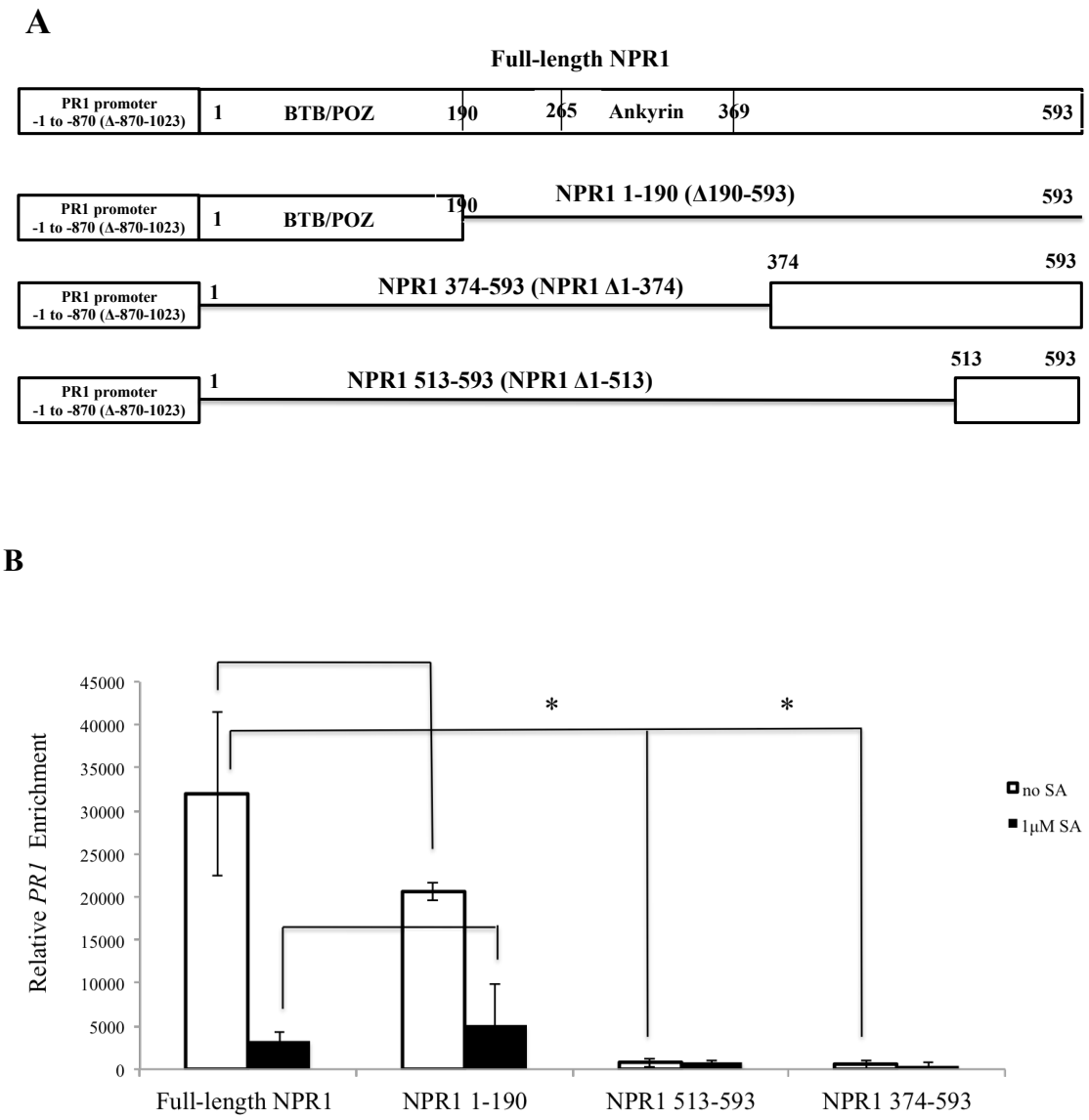


Figure 10. Possible location of the DNA binding domain of NPR1 for binding to the -870 fragment of the PR1 promoter in the absence of SA

A. Schematic representation of NPR1 deletion constructs. The first box from the top shows full-length NPR1 (amino acids 1-593). The second box indicates NPR1-190 (Δ190-593) that amino acids 190 to 593 deleted. The third box represents NPR1 374-593 (Δ1-374), which amino acids 1 to 374 deleted. Last box shows NPR1-Δ513 (Δ1-513) that amino acids 1 to 513 deleted. The white box shows the -1 to -870 fragment of the PR1 promoter that was used as a DNA source, along with NPR1 deletion constructs employed in ChIP experiments.
B. Histogram illustrating modified ChIP experiments performed on both C and N terminus deletion constructs of NPR1 for binding to the -870 fragment of the PR1 promoter, before and after addition of SA, confirming the possible localization of the NPR1 DNA binding domain between amino acids 1 and 190. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment amounts were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

In order to further narrow down the location of the NPR1 binding site domain for binding to the -870 fragment of the PR1 promoter, a new series of experiments were performed with C-terminus deletion construct NPR1 110-593 (NPR1 Δ1-110) before and after addition of SA (Figure 11A). NPR1 Δ1-110 lacks the first 110 N-terminal amino acids. For this series of experiments, NPR1 1-190 (Δ190-593) and full-length constructs were employed as positive controls.

Remarkably, NPR1 110-593 (NPR1 Δ1-110) showed similar amounts of relative *PR1* enrichment compared to full-length NPR1, both in the presence and absence of SA (Student two-tailed t-test, p-values are 0.8947, and 0.0618 respectively; p-value < 0.05), with larger *PR1* enrichment values observed in the absence of SA. Consistent with previous results (Figure 10B), a similar amount of relative *PR1* enrichment was observed for NPR1 1-190 (Δ190-593) compared with full-length NPR1, both in the presence and absence of SA (Student two-tailed t-test, p-values are 0.6932, and 0.2390 respectively; p-value < 0.05), with larger enrichment values observed in the absence of SA (Figure 10B).

These findings demonstrated that in the absence of SA, the first 110 amino acids in the 1-190 sequence located in the N-terminus region did not play a significant role in the binding of NPR1 to the -870 fragment of the PR1 promoter. Additionally, a comparison of relative *PR1* enrichment values of full-length NPR1, NPR1 1-190 (Δ190-593), and NPR1 110-593 (NPR1 Δ1-110) clearly confirmed that the DNA binding domain of NPR1 for binding to the -870 fragment in the absence of SA is localized between amino acids 110-190 in the N-terminus region.

Following, investigations were conducted to determine if the binding of NPR1 110-593 (NPR1 Δ1-110) amino acid sequence to the -870 fragment in the absence of SA is facilitated by a metal co-factor. For this investigation,

Ethylenediaminetetraacetic acid (EDTA) was used as a metal chelating agent in a series of modified ChIP experiments employed to analyze NPR1 1-190 (Δ190-593) binding to the -807 fragment of the PR1 promoter before and after addition of 1μMSA (as positive control) and 10mM EDTA.

Our findings revealed that in the presence of EDTA (metal chelation), relative *PR1* enrichment for NPR1 1-190 (Δ190-593) was completely abolished, with statistically significant differences found in comparison to values obtained before addition of EDTA (Student two-tailed t-test, p-value is 0.0008; p-value < 0.05) (Figure 11C). Consistent with our previous results (Figure 11B and 10B), relative *PR1* enrichment values were found to significantly decrease after addition of SA as compared to results obtained prior to SA addition (Student two-tailed t-test, p-value is 0.0408; p-value < 0.05) (Figure 11C). Chiefly, the obtained results confirmed that the binding of NPR1 1-190 (Δ190-593) domain amino acid sequence 110-190 to the -870 fragment of the PR1 promoter occurs in the absence of SA through a metal co-factor.

Our findings regarding the role of a metal co-factor in the binding of NPR1 to the -870 fragment (Figure 11C) led to further exploration on the exact identity of the amino acid associated in this process. As shown by Cao et al. (1997), the *npr1-2* allele mutation exhibits only a minor induction of *PR1* expression after SA treatment. The *npr1-2* mutation converts cysteine residue number 150 to tyrosine 150. The importance of cysteine residue number 150 in NPR1 plant immune response activity allows for the hypothesis that cysteine residue number 150 could be the amino acid under investigation. To test this hypothesis, NPR1 1-190 (Δ190-593) was submitted to a mutation in its Cysteine residue 150 to tyrosine 150, and a series of modified ChIPs experiments were conducted for binding of NPR1 1-190 (Δ190-593) C150Y to the -870 fragment of the PR1 promoter. For these experiments, NPR1 1-190 (Δ190-

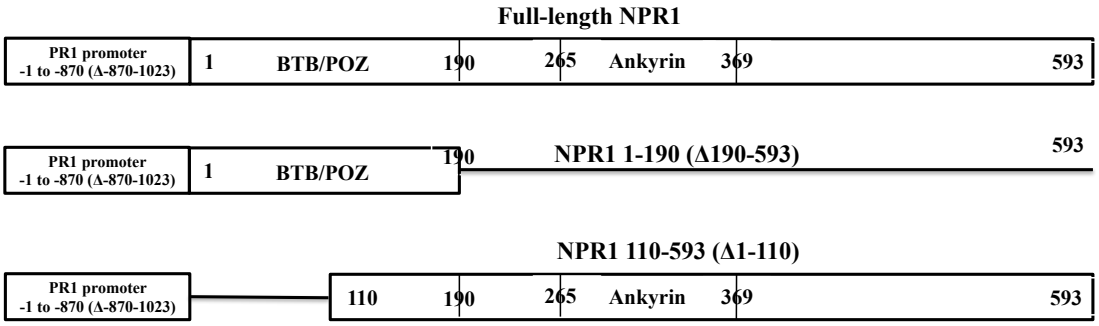
593) was employed as a positive control measure prior to and after treatment with SA and EDTA.

The obtained results demonstrated that relative *PR1* enrichment for NPR1 1-190 (Δ190-593) C150Y was abolished completely, with a significant observed difference compared to the control NPR1 1-190 (Δ190-593) (Student two-tailed t-test, p-value is 0.0324; p-value < 0.05) (Figure 11D). Consistent with our previous results, relative *PR1* enrichment for NPR1 1-190 (Δ190-593) was abolished after addition of EDTA (Student two-tailed t-test, p-value is 0.0008; p-value < 0.05) (Figure 11D). As well, a significant reduction in relative *PR1* enrichment was observed (Figure 10B, 11B, 11C) after addition of SA as compared to results obtained prior to SA addition (Student two-tailed t-test, p-value is 0.0408; p-value < 0.05) (Figure 11D).

These findings further confirmed that in the absence of SA, the binding of NPR1 to the -870 fragment of the PR1 promoter is a metal-dependent process; in addition, the amino acid cysteine number 150 located in the N-terminus region of NPR1 was shown to be critical to the transition-metal binding activity of NPR1 to the PR1 promoter.

In sum, the obtained results further confirm that the DNA binding domain of NPR1 for binding to the -870 fragment of the PR1 promoter is located in the N-terminus region, between amino acids 110 to 190. Further, in the absence of SA, the NPR1 binding has been revealed to occur through a metal co-factor and Cys^{150} shown to be critical for metal interaction.

A



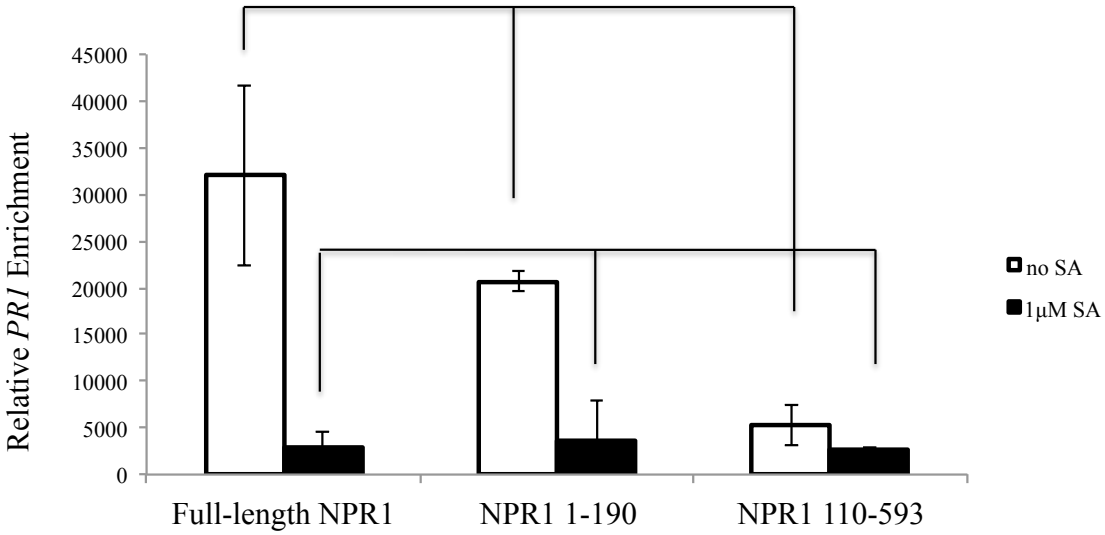
110 **150**

KLELKEIAKDYEVGFDSVVTVLAYVYSSRVRPPPKGVSECADENCCHVACRPAVDFMLEVLYLAFIFKIP

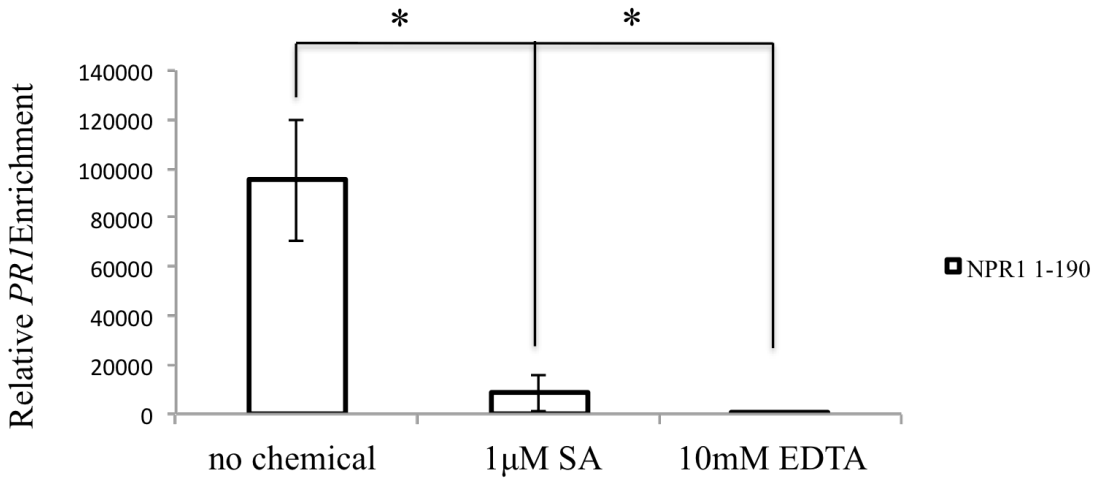
ELITLYQRHLLDVVDKVVIEDTLVILKLANICGKACMKLLDRCKEIIVKSNVDMVSLEKS

190

B



C



D

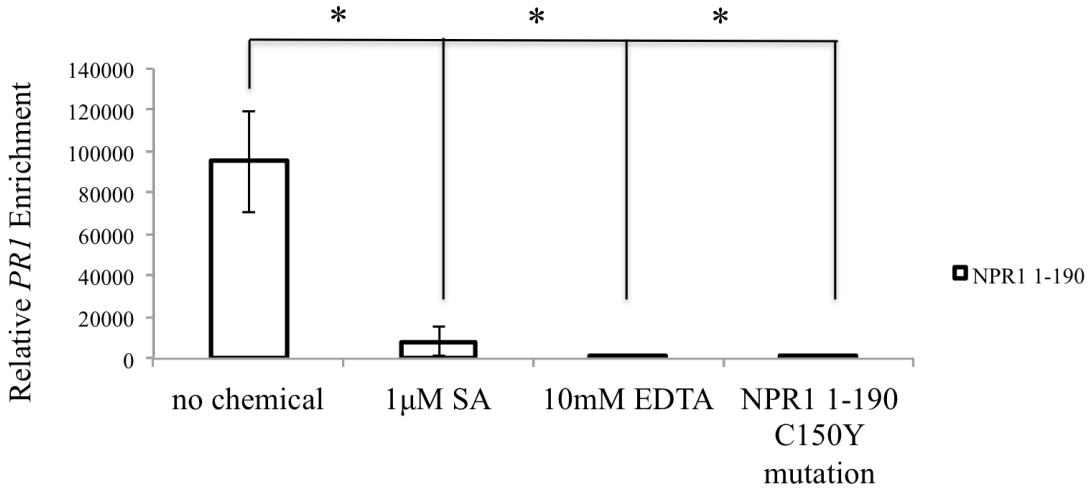


Figure 11. DNA binding domain of NPR1 for binding to the -870 fragment of the PR1 promoter, located between amino acids 110 to 190. Binding occurs in the absence of SA through Cys^{150} via transition metal

A. Schematic representation of NPR1 deletion constructs. The first box from the top shows full-length NPR1 (amino acid 1-593). The second box indicates NPR1 1-190 (Δ190-593) that amino acids 190 to 593 deleted. The third box represents NPR1 110-593 (Δ1-110), which amino acids 1 to 110 deleted. White box shows the -1 to -870 fragment of the PR1 promoter used as DNA source along with NPR1 deletion constructs employed in ChIP experiments. Sequence of NPR1 N-terminus region amino acid 110-190 shows Cys^{150} position.
B. Histogram illustrating modified ChIP experiments performed on the NPR1 N-terminus deletion constructs and the -870 fragment of the PR1 promoter before and after addition of SA, confirming that the first 110 amino acids of the N-terminus are not associated with NPR1 binding to -870 fragment of the PR1 promoter. ChIPs were performed with anti-

Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

C. Histogram illustrating modified ChIP experiments performed on NPR1 1-190 (Δ190-593) and the -870 fragment of the PR1 promoter after EDTA addition, confirming that the binding of NPR1 to the PR1 promoter -870 fragment is metal dependent. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).
D. Histogram illustrating modified ChIP experiments performed on NPR1 1-190 (Δ190-593), NPR1 1-190 (Δ190-593) C150Y mutation, and the -870 fragment of the PR1 promoter after SA addition, confirming that that the binding of NPR1 to the -870 fragment is metal dependent and that Cys^{150} is a critical factor in this metal interaction. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

4.9 Exclusive binding of NPR1 protein to circular double stranded DNA

The successful determination of NPR1 binding sites on the PR1 promoter in the presence and absence of SA (Figure 4 and 5), as well as the subsequent clarifications regarding the localization of the DNA binding domain positions on NPR1 (Figure 8 and 11) led to further specification inquiries in regards to the role DNA shape plays in NPR1 binding. Concisely, the aim of the next experiments centered on whether NPR1 favors single stranded versus double stranded DNA, or whether it binds linear versus circular-form DNA. Since the PR1 promoter sequence was cloned to a plasmid, different tests could be performed, including restriction digestion and boiling.

First, NPR1 binding was tested to both the -701 and -870 fragments of the PR1 promoter (Figure 12A) after digestion of both fragments with two different restriction enzymes, including *Bgl*I and *SnaBI*. Sequence specific recognition sites for both enzymes were located outside of the region where NPR1 binds to the -701 and -870 fragments of the PR1 promoter. Next, -701 and -870 fragment samples were boiled in order to denature the double stranded form of the plasmid to single stranded form, followed by testing of their bindings to NPR1 protein.

The obtained results revealed that relative *PR1* enrichment was severely reduced when both -701 and -870 fragments of the PR1 promoter were digested with the *SnaBI* enzyme and denatured after boiling, as well as upon submission to a combination of both digestion and boiling (Figure 12B). The observed reductions in relative *PR1* enrichment for all conditions were significant compared to *PR1* enrichment values obtained for both the -701 and -870 fragments in circular and double stranded forms. Student two-tailed t-test p-values for the -870 fragment submitted to i) digestion, ii) boiling, and iii) combinations of both digestion and

boiling, were 0.0015, 0.0266, and 0.0165, respectively (p-value < 0.05). Similarly, student two-tailed t-test p-values for the -701 fragment submitted to i) digestion, ii) boiling, and iii) combinations of both digestion and boiling, were 0.0731, 0.0723, and 0.0556, respectively (p-value < 0.05).

The NPR1 recruitment to the -701 and -870 fragments of the PR1 promoter was also tested after samples were digested with *Bgl*I enzyme (Figure S9) prior to and after addition of SA. As expected, similar to *SnaBI* digestion results (Figure 12B), relative *PR1* enrichment was completely abolished before and after addition of SA (Figure 12C).

Given that NPR1 only binds to circular and double stranded forms of DNA, and given that a rich AT sequence is present around the -701 fragment region of the PR1 promoter, further investigations were then conducted regarding the existence of secondary DNA structure formations that resulted in stabilization of binding between NPR1 and the PR1 promoter. For this purpose, Mfold DNA secondary structure prediction software (Zuker, 2003) was employed to predict all possible formations, including hairpins, stem-loops, and cruciforms, capable of occurring in both the -701 and -870 fragments regions (Figure S3). Prediction analysis prompted the hypothesis that the formation of secondary structures around these specific NPR1 binding sites perhaps stabilizes the interactions occurring between NPR1 and the PR1 promoter. Unfortunately, a demonstration of the formation of secondary structures in *vitro* using the classical *S1* endonuclease enzyme mapping approach was not possible (data not shown). As such, the hypothesis remains to be confirmed.

In conclusion, our findings demonstrate that the NPR1 protein only binds to double-stranded and circular forms of the -701 and -870 fragments of the PR1 promoter

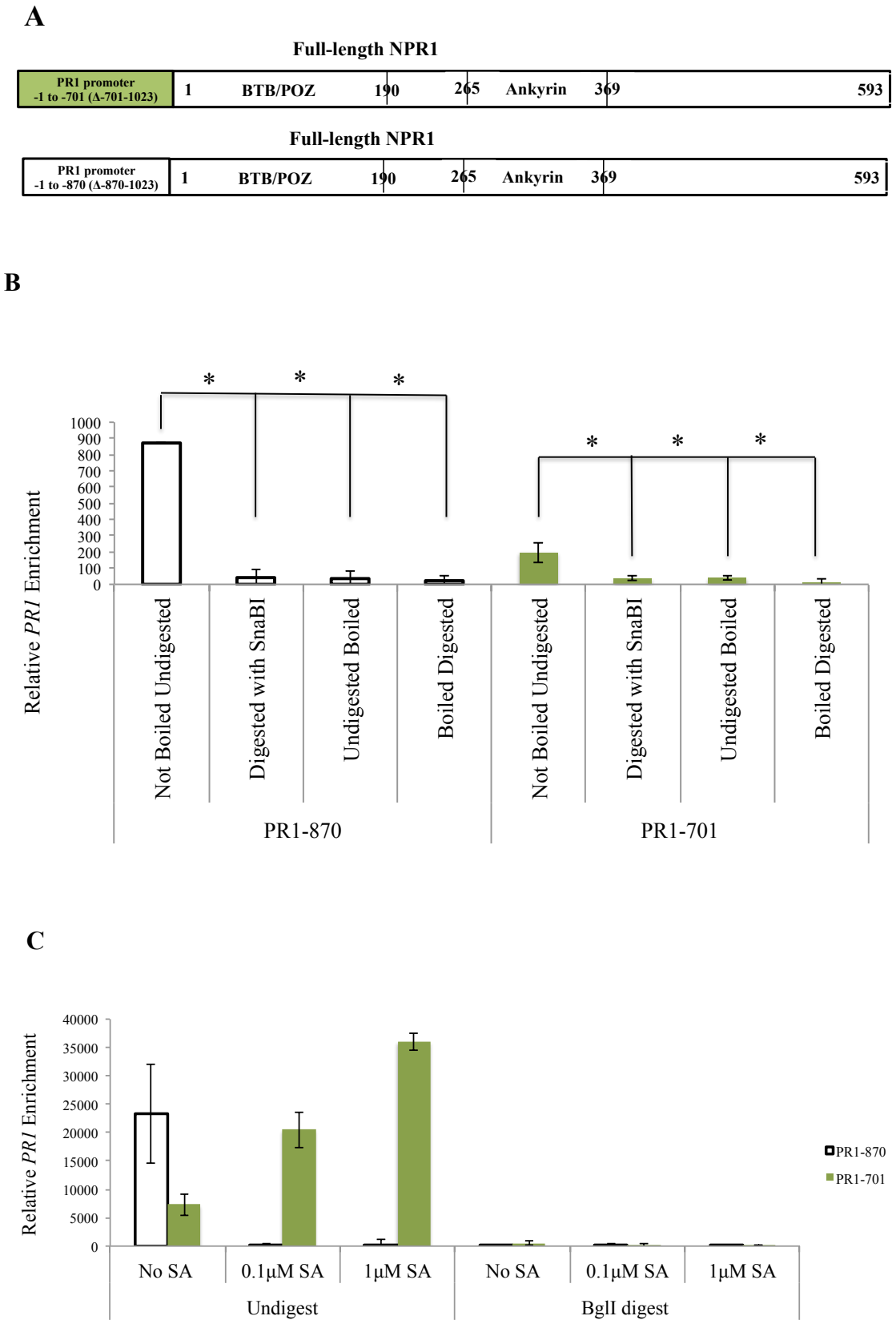


Figure 12. Binding of NPR1 to the -701 and -870 fragments of the PR1 promoter

according to DNA form

A. Schematic representation of PR1 promoter deletion constructs and full-length NPR1 protein (amino acids 1-593) used for modified ChIP experiments. Fragments of the PR1 promoter sequence cloned into a plasmid and used as DNA source in modified ChIP experiments. The white box indicates -1 to -870 fragment of the PR1 promoter sequence. The green box shows the -1 to -701 fragment of the PR1 promoter sequence.
B. Histogram illustrating modified ChIP experiments performed on the -701 and -870 fragments of the PR1 promoter, confirming that NPR1 only binds to circular and double-stranded forms of PR1 promoter fragments. Restriction digestions were conducted with the use of *SnaBI* enzyme. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).
C. Histogram illustrating modified ChIP experiments performed on the -701 and -870 fragments of the PR1 promoter before and after addition of SA, further confirming that NPR1 only binds circular and double stranded forms of PR1 promoter fragments. Restriction digestions were conducted with the use of *Bgl*I enzyme. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates.

4.10 In vivo and in vitro analyses of NPR1-PR1 promoter binding inhibition by 4-hydroxy benzoic acid

As previously discussed in section (4.3), an abolishment of relative *PR1* enrichment was observed for the -701 and -870 fragments of the PR1 promoter upon addition of 0.1 and 1μM of 4-OH-BA (Figure 3D). Thus, SA was further concluded to positively regulate the binding of NPR1 to the -701 fragment of the PR1 promoter, while 4-OH-BA NPR1 was demonstrated to negatively affect the binding of NPR1 to the PR1 promoter. In related work conducted by Wu et al. (2012), 4-OH-BA has also been shown to compete with SA for NPR1 binding however, 4-OH-BA was shown to forgo binding to NPR1.

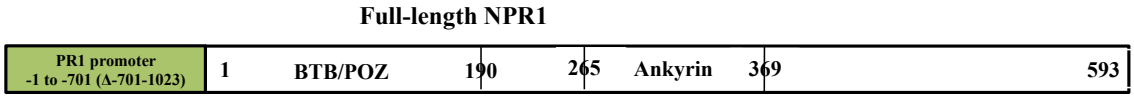
These results led to the hypothesis that 4-OH-BA might act as a DNA binding inhibitor by competing with SA for NPR1 binding to the PR1 promoter. To test this hypothesis, modified ChIP experiments were conducted for NPR1 binding to the -701 fragment of the PR1 promoter (Figure 13A) prior to and after addition of 1 and 10μM SA, as well as prior to and after addition of combinations of 1 and 10μM SA with 1, 10, and 80μM 4-OH-BA (Table 3.1) (Figure 13B). The combined NPR1 and -701 fragment of the PR1 promoter were incubated with 4-OH-BA (1, 10, and 80μM) for 15 min, followed by subsequent addition of either 1 or 10μM SA. Samples were then left to stand for another 15 min to allow cross-linking to occur.

The obtained results revealed that when 1μM of SA was added to 1, 10, and 80μM of 4-OH-BA, the binding of NPR1 to the -701 fragment of the PR1 promoter was highly affected, with a significant reduction in relative *PR1* enrichment values as compared to when only 1μM SA was added (Figure 13C) (Student two-tailed t-test, p-values are 0.0491, 0.0285, and 0.0325 respectively; p-value < 0.05). Similarly, upon addition of 10μM of SA to 1, 10, and 80μM of 4-OH-BA, the binding of NPR1 to the

-701 fragment of the PR1 promoter was revealed to be affected, as the amount of relative *PR1* enrichment was significantly reduced as compared to results obtained for single addition of 10μM SA (Figure 13D). Student two-tailed t-test analyses produced p-values of 0.0027, 0.0111, and 0.0043 for 1, 10, and 80μM of 4-OH-BA combined with 10μM SA, respectively, while single addition of 10μM SA produced a p-value < 0.05.

These findings suggest that 4-OH-BA competes with SA, negatively regulating NPR1 binding to the PR1 promoter. However, no drastic changes were observed between the 1, 10, and 80μM concentrations of 4-OH-BA. Presumably, the binding sites on NPR1 were already saturated at lower concentrations of 4-OH-BA; consequently, increases in 4-OH-BA concentration did not contribute to further drastic changes.

A

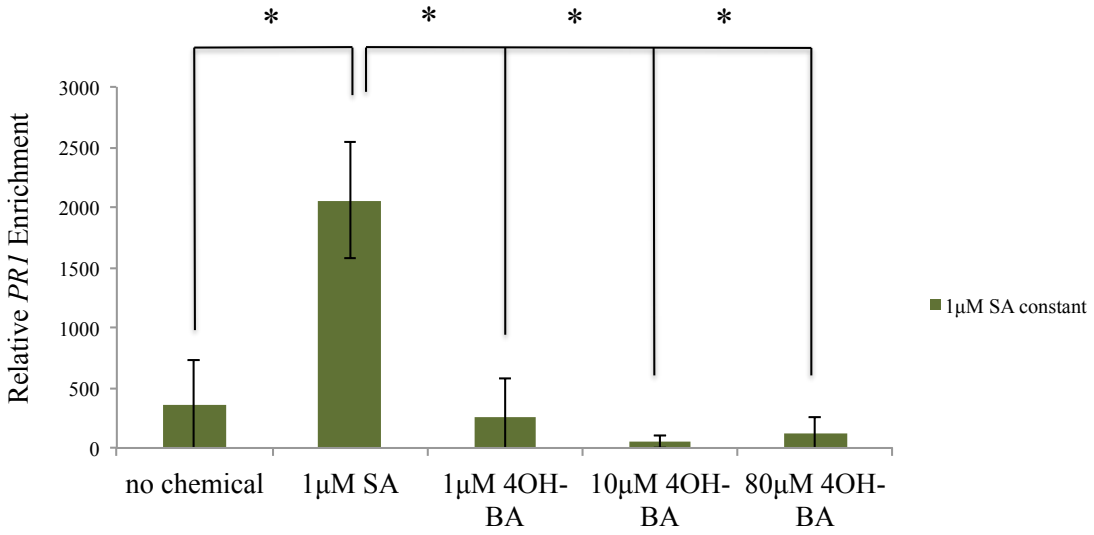


B

Salicylic acid (SA)

4 Hydroxy Benzoic acid (4-OH-BA)

C



D

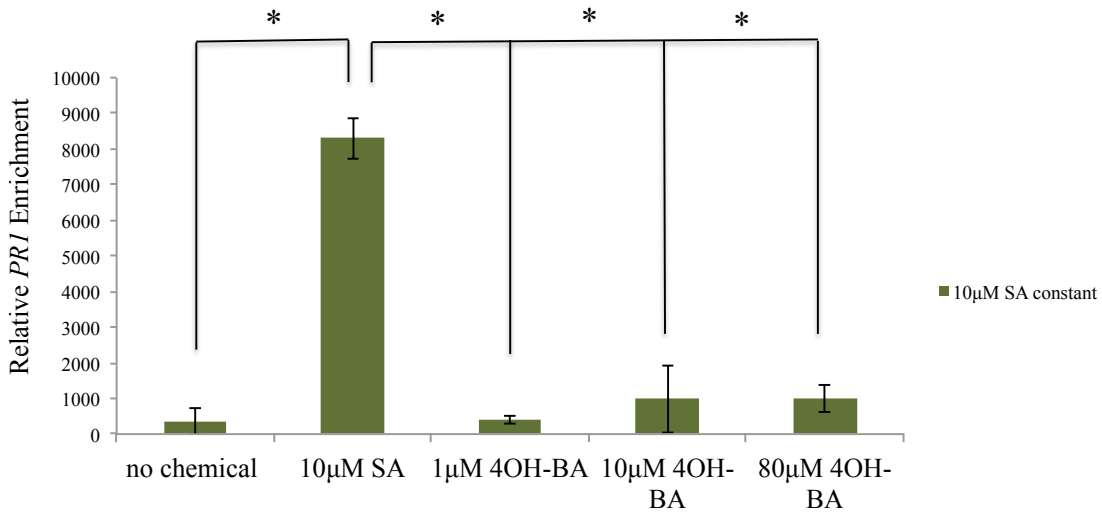


Figure 13. Binding site competition between 4-OH-BA and SA, and negative regulation of 4-OH-BA in binding of NPR1 to the -701 fragment of the PR1 promoter

A. Schematic representation of PR1 promoter and full-length NPR1 protein (amino acids 1-593) used for modified ChIP experiments. The green box shows the -1 to -701bp fragment of the PR1 promoter sequence.
B. Structures of SA and 4-OH-BA.
C. Histogram illustrating modified ChIP experiments performed on NPR1 and the -701 fragment of the PR1 promoter after addition of SA and 4-OH-BA. First, 1, 10, and 80µM of 4-OH-BA were added, followed by addition of 1µM SA. 4-OH-BA competed with SA and negatively affected NPR1 binding. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).
D. Histogram illustrating modified ChIP experiments performed on NPR1 and the -701 fragment of the PR1 promoter after addition of SA and 4-OH-BA. First 1, 10, and 80µM of 4-OH-BA were added, followed by addition of 10µM SA. 4-OH-BA competed with SA and negatively affected NPR1 binding. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

Following these initial results, further research was undertaken to explicate the competitive relationship between 4-OH-BA and SA. For this, relative *PR1* enrichment amounts were obtained for binding of two NPR1 C-terminus deletion constructs, NPR1 513-593 (NPR1 Δ1-513) and NPR1 533-593 (NPR1 Δ1-533) to the - 701 fragment of the PR1 promoter (Figure 14A) under the presence of a series of low-to-high concentrations of 4-OH-BA (1, 0.1, 0.00001, 0.000001, and 0.0000001μM), and 1μM SA. It is important to note that NPR1 513-593 (NPR1 Δ1-513) and NPR1 533-593 (NPR1 Δ1-533) were particularly chosen for these experiments for two distinct reasons. Foremost, NPR1 513-593 (NPR1 Δ1-513) has been shown to contain the specific sites (Cysteine 521 and 529 residues) on NPR1 for SA perception (Wu et al., 2012). Additionally, as found in section (4.7), our results have revealed that in the presence of SA, the DNA binding domain of NPR1 for binding to the -701 fragment of the PR1 promoter is also localized between amino acids 513 to 535 in the C-terminus region of NPR1 (Figure 8). Also, Rochon et al., (2006) showed that amino acids 513 to 533 are critical for transactivation function of NPR1 in presence of SA.

As can be seen in Figure 14B, a negative correlation between 4-OH-BA and relative *PR1* enrichment amounts was found to exist for the NPR1 513-593 (NPR1 Δ1-513); in other words, higher concentrations of 4-OH-BA yielded lower relative *PR1* enrichment values. Consistent with previous results for NPR1 513-593 (NPR1 Δ1-513) (Figures 6,7, and 8), no significant differences in relative *PR1* enrichment were observed between values obtained before and after addition of SA (Figure 14B) (Student two-tailed t-test, p-value is 0.7235; p-value < 0.05).

Student two-tailed t-tests showed that significantly lower amounts of relative *PR1* enrichment were obtained after addition of 1 and 0.01μM 4-OH-BA (p-values were 0.0596 and 0.0264 respectively; p-value < 0.05) as compared to the amount of

PR1 enrichment obtained after addition of 1μM SA (Figure 14B). As well, lower relative *PR1* enrichment values were observed after addition of 0.00001, 0.000001, and 0.0000001μM 4-OH-BA as compared to the amount of relative *PR1* enrichment obtained after addition of 1μM SA. However, a student two-tailed t-test analysis showed that the *PR1* enrichment values obtained after addition of 0.00001, 0.000001, and 0.0000001μM 4-OH-BA were not statistically significant compared to *PR1* enrichment values obtained after addition of 1μM SA (p-values were 0.0949, 0.1239, and 0.6085 respectively; p-value < 0.05) (Figure 14B).

Similar to NPR1 513-593 (NPR1 Δ1-513) results (Figure 14B), an inverse relationship was observed between 4-OH-BA concentrations values (1 to 0.0000001μM) and the amounts of relative *PR1* enrichment obtained for NPR1 533-593 (NPR1 Δ1-533) (Figure 14C). No significant differences were noted in regards to relative *PR1* enrichment values obtained for NPR1 533-593 (NPR1 Δ1-533) before and after addition of SA (Figure 14C) (Student two-tailed t-test, p-value is 0.5087; p-value < 0.05).

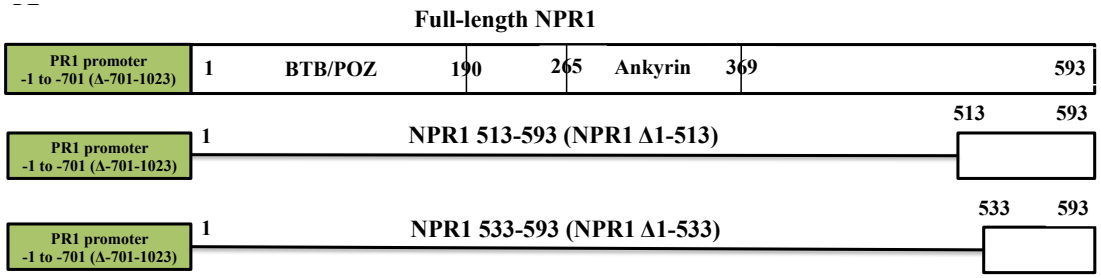
However, significantly lower amounts of relative *PR1* enrichment were obtained after addition of 1, 0.1, and 0.00001μM 4-OH-BA as compared to *PR1* enrichment values obtained after addition of 1μM SA (Figure 14C) (Student two-tailed t-test, p-values are 0.0480, 0.0435, and 0.0548 respectively; p-value < 0.05). As well, no significant differences were noted in regards to relative *PR1* enrichment values obtained after addition of 0.000001 and 0.0000001μM 4-OH-BA as compared to *PR1* enrichment values obtained after addition of 1μM SA (Figure 14C) (Student two-tailed t-test, p-values are 0.1113, and 0.2721 respectively; p-value < 0.05).

For NPR1-Δ513, after addition of 0.00001μM 4-OH-BA, the obtained relative *PR1* enrichment was similar to enrichment values obtained after addition of 1μM SA

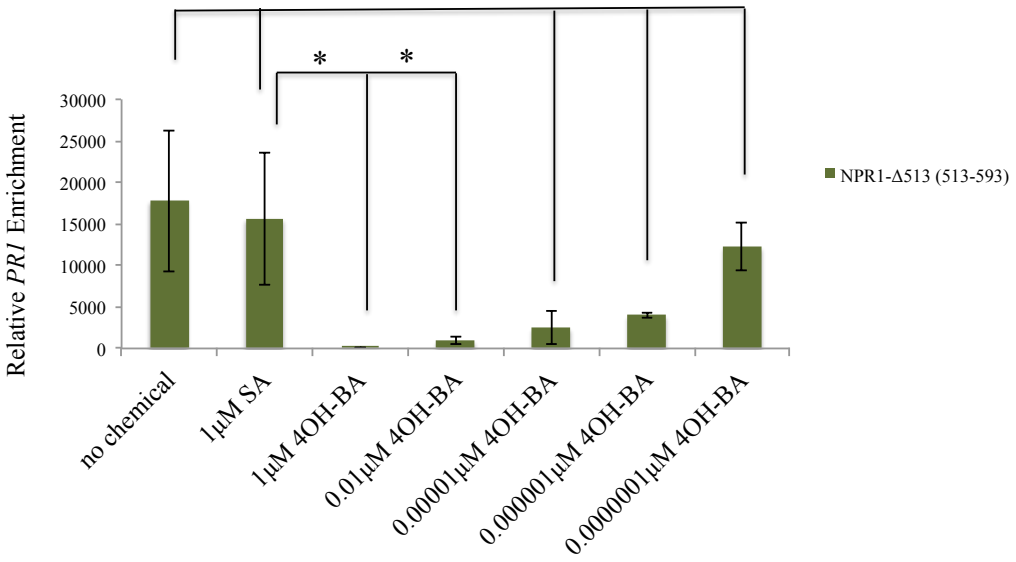
(Figure 14B). Conversely for NPR1 533-593 (NPR1 Δ1-533), enrichment values similar to those obtained after addition of 1μM SA were observed after addition of 0.000001μM 4-OH-BA (Figure 14C). These observations served to confirm that 4-OH-BA impacts the relative enrichment of *PR1* for NPR1 513-593 (NPR1 Δ1-513) to a considerably faster degree in comparison to NPR1 533-593 (NPR1 Δ1-533).

The obtained results for this series of experiments suggest that 4-OH-BA inhibits the binding of NPR1 to the -701 fragment of the PR1 promoter via competition with SA for NPR1 binding sites. Also, the faster degree of change observed in *PR1* enrichment amounts for NPR1 513-593 (NPR1 Δ1-513) as compared to NPR1 533-593 (NPR1 Δ1-533) further confirms that 4-OH-BA acts as a potent inhibitor of DNA-binding activity of NPR1, even at low concentration levels.

A


Full-length NPR1
PR1 promoter
-1 to -701 (Δ-701-1023)
1
BTB/POZ
190
265
Ankyrin
369
593
513
593
NPR1 513-593 (NPR1 Δ1-513)
533
593
NPR1 533-593 (NPR1 Δ1-533)

B


Relative PR1 Enrichment
30000
25000
20000
15000
10000
5000
0
*
*
no chemical
1μM SA
1μM 4OH-BA
0.01μM 4OH-BA
0.00001μM 4OH-BA
0.000001μM 4OH-BA
0.0000001μM 4OH-BA
NPR1-Δ513 (513-593)

C

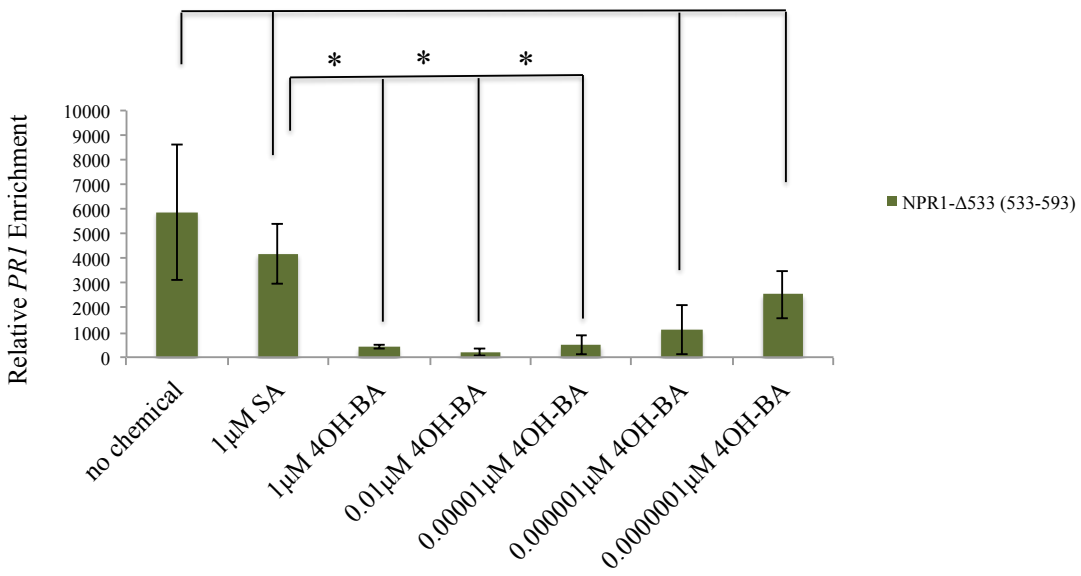

Relative PR1 Enrichment
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
*
*
*
no chemical
1μM SA
1μM 4OH-BA
0.01μM 4OH-BA
0.00001μM 4OH-BA
0.000001μM 4OH-BA
0.0000001μM 4OH-BA
NPR1-Δ533 (533-593)

Figure 14. 4-OH-BA inhibition of binding of NPR1 to the -701 fragment of the PR1 promoter

A. Schematic representation of the PR1 promoter and NPR1 protein deletions constructs used for modified in vitro ChIP experiments. Green box shows the -1 to -701bp fragment of the PR1 promoter sequence. The top box indicates Full-length NPR1 protein (amino acid 1 to 593). The second box shows NPR1 513-593 (NPR1 Δ1-513), which amino acids 1 to 513 deleted. The third box represents NPR1 533-593 (NPR1 Δ1-533) that amino acids 1 to 533 deleted.
B. Histogram illustrating modified ChIP experiments performed on NPR1 513-593 (NPR1 Δ1-513) and the -701 fragment of the PR1 promoter before and after addition of 1μM SA, and 1, 0.1, 0.00001, 0.000001, 0.0000001μM 4-OH-BA, confirming that 4-OH-BA inhibits the binding of NPR1 513-593 (NPR1 Δ1-513) to the -701 fragment. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).
C. Histogram illustrating modified ChIP experiments performed on NPR1 533-593 (NPR1 Δ1-533) and -701 fragment of the PR1 promoter before and after addition of 1μM SA, and 1, 0.1, 0.00001, 0.000001, 0.0000001μM 4-OH-BA, confirming that 4-OH-BA inhibits the binding of NPR1 533-593 (NPR1 Δ1-533) to the -701 fragment. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05)

After successful in vitro determination of 4-OH-BA as a potent inhibitor of NPR1 binding to the PR1 promoter (Figure 13 and Figure 14), an in vivo experimental set-up was established to further test the potential role of 4-OH-BA in live plants. In this set-up, plants were sprayed with 4-OH-BA and SA with the aim to test the regulation of *PR1* transcription that occurs as a result of the antagonist relationship between 4-OH-BA and SA. Quantitative reverse transcriptase PCR (qRT-PCR) analyses in wild type (WT) *Arabidopsis* (Figure 15A) were performed prior to treatment with SA and 4-OH-BA, 12 hours past initial 300μM SA treatment, and 12 hours following combination treatments composed of 300μM SA and varying concentrations of 4-OH-BA (30, 3, and 0.3μM). After RNA isolation from untreated and treated *Arabidopsis* leaves and its subsequent conversion to cDNA, *PR1* transcripts were quantified through the use of primers specific to the *PR1* gene.

Our findings indicated that PR1 transcripts were more abundant in SA-induced samples in comparison to non-induced samples (Figure 15B). As well, *PR1* transcripts were observed in more abundance when plants were sprayed with a combination of 300μM SA and 0.3μM 4-OH-BA. However, *PR1* transcripts were less substantial for plants sprayed with a combination of 300 μM SA and either 3 or 30μM 4-OH-BA (Figure 15B). Together, the obtained results suggest that 4-OH-BA negatively regulated *PR1* transcription as opposed to SA at higher concentrations (3 or 30μM), while at lower concentrations (0.3μM), *PR1* transcription was able to return back to SA-induced levels.

Next, we approached the possibility of 4-OH-BA analogues also negatively regulating *PR1* transcripts via SA competition. In related work, Wu et al. (2012) demonstrated that other analogs of SA do not bind to NPR1; in comparison to SA, they have been shown to be poor inducers of *PR1* transcripts. Bearing this in mind, 4-

OH-BA analogues were analyzed in a manner similar to the previously described in vivo approach undertaken for 4-OH-BA. For this series of experiments, qRT-PCR analyses were performed to determine the abundance of *PR1* transcripts found in wild-type *Arabidopsis* leaves with the following experimental parameters: i) non-induced leaves, ii) 12 hours past 300μMSA induction, iii) 12 hours after induction of combinations of both 300μMSA and 3-OH-BA (30, 3, and 0.3μM), INA (30, 3, and 0.3μM), catechol (30, 3, and 0.3μM), and Benzoic acid (BA) (30, 3, and 0.3μM) (Table 3.2) (Figure 15A and Figure S4A).

Our findings revealed that *PR1* transcript levels in plants sprayed with a combination of SA and 3-OH-BA (Figure S4B), SA and INA (Figure S4D), and SA and BA (Figure S4C) were similar to those obtained for plants treated with SA and 4-OH-BA (Figure 15B), although this was only true for SA-induced leaves. Conversely, a combination of SA with different concentrations of either 3-OH-BA, INA, and BA showed inconsistent *PR1* transcripts, with an overall decrease in *PR1* transcript levels for all concentrations of 3-OH-BA and BA (30, 3, and 0.3μM) when applied in conjunction with SA (Figure S4B and C).

Interestingly, opposed to 4-OH-BA, 3-OH-BA, and BA results (Figure 15B, Figure S4B and C) high *PR1* transcripts levels similar to those obtained for SA-induced leaves (Figure 15C), were found for plants 12 hours past the initial spraying with a combination of 300μM SA and catechol (30, 3, and 0.3μM) (Figure 15A), suggesting that catechol competed with SA synergistically, dissimilar to how 4-OH-BA competes for NPR1 binding to the PR1 promoter, and ultimately, *PR1* expression.

Next, qRT-PCR analyses were performed to quantify the abundance of *PR1* transcripts in wild-type and *PAL*-mutant *Arabidopsis* plants before and after treatment for 12 hours with 300μM SA. The Phenylalanine Ammonia Lyase (PAL) gene has

been shown to encode for the *PAL* enzyme that acts as an intermediate in the 4-OH-BA synthesis pathway (Kohle et al., 2002). Based on our previous results, higher or similar amounts of *PR1* transcripts were expected for *PAL* mutant plants compared to the SA-induced wild-type plants. As expected, *PR1* transcripts levels were higher for *PAL* mutant plants in comparison to values obtained for SA-induced wild type plants (Figure S4E). However, *PAL1* and *PAL2 (*two different alleles of *PAL* mutant) yielded differing *PR1* transcripts levels; as such, the results were considered to be inconclusive.

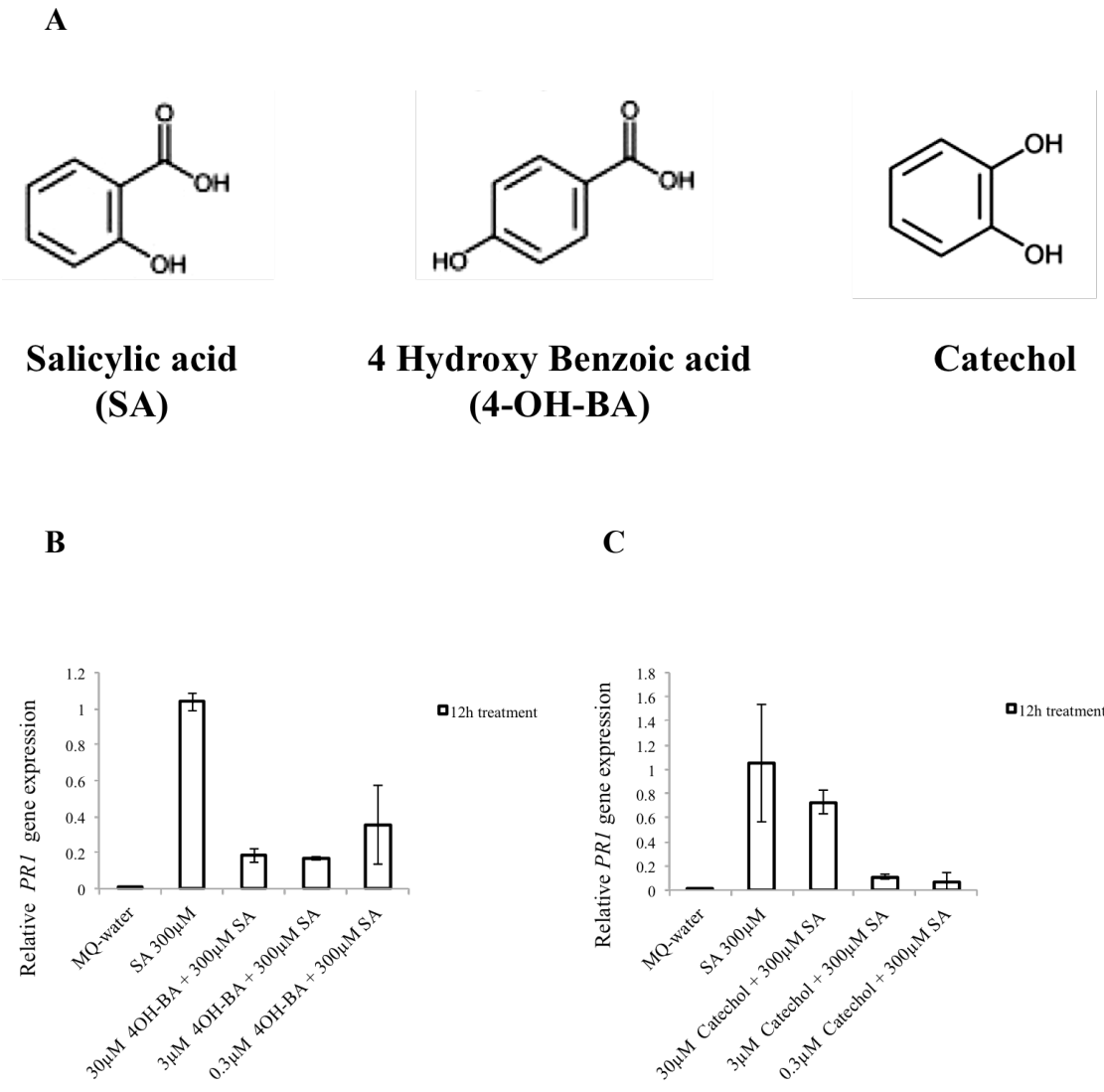


Figure 15. Competition of 4-OH-BA and catechol with SA for *PR1* expression regulation

A. Structures of SA and its inactive analogues, 4-OH-BA and catechol.
B. Histogram illustrating the expression profile of the *PR1* gene prior to SA treatment, 12 hours after 300µM SA treatment, and 12 hours after a combination of 300µM SA with 30, 3, and 0.3µM 4-OH-BA. Quantitative reverse transcriptase PCR (RT-PCR) was conducted in 6 weeks old *Arabidopsis thaliana* leaves. Expression was normalized to the expression of the constitutive gene *ubiquitin 5*.
C. Histogram illustrating the expression profile of the *PR1* gene prior to SA treatment, 12 hours after 300µM SA treatment, and 12 hours after a combination of 300µM SA with 30, 3, and 0.3µM catechol. Quantitative reverse transcriptase PCR (RT-PCR) was conducted in 6-week old *Arabidopsis thaliana* leaves. Expression was normalized to the expression of the constitutive gene *ubiquitin 5*.

4.11 Binding of NPR2, NPR3, and NPR4 proteins to the PR1 promoter

Five other *NPR1*-like genes have been identified in the *Arabidopsis* genome, namely *NPR2*, *NPR3, NPR4, NPR5* (*BLADE-ON-PETIOLE2*), and *NPR6* (*BLADE-ON-PETIOLE1*) (Hepworth et al., 2005; Liu et al., 2005). NPR2 shares the highest sequence similarity with NPR1 (61.89%) followed by NPR3 (36.67%), NPR4 (36.24%), NPR5 (20.16%), and NPR6 (21.63%) (Table S1 and Figure S5A).

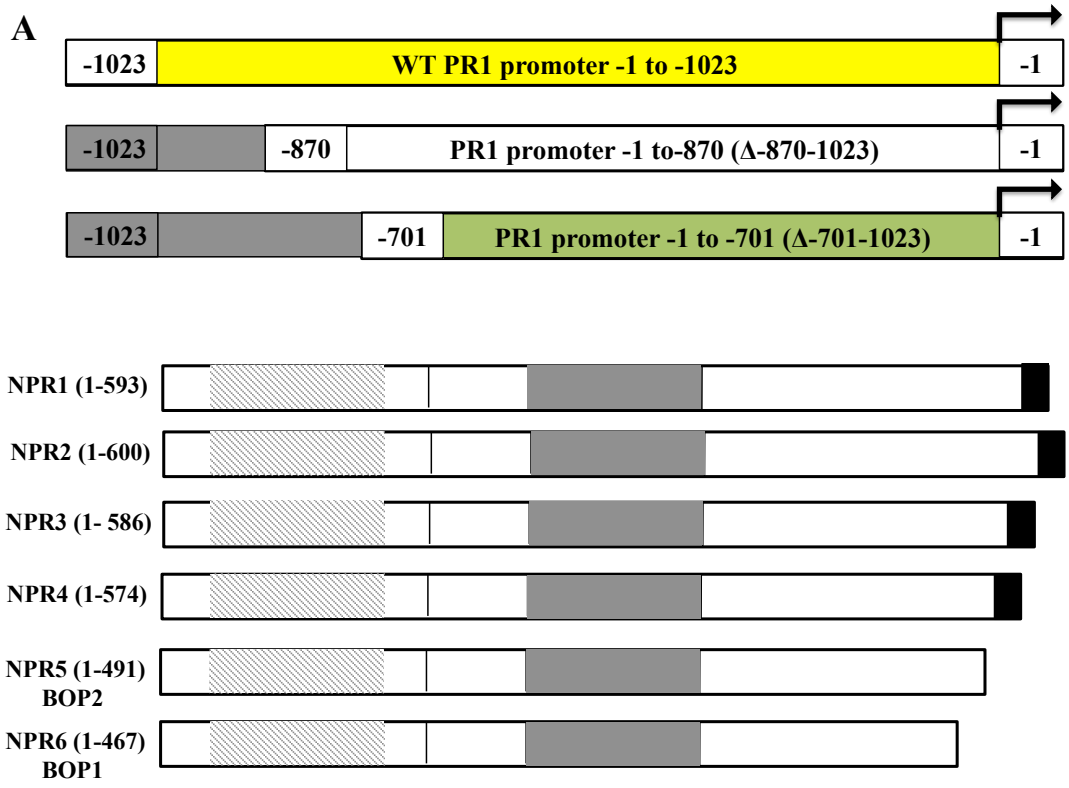
Although NPR3 and NPR4 have been shown to be involved in plant defence mechanisms, contradictory results from different groups have impeded a clearer, more concise understanding of the role these two genes play in plant resistance responses. As such, we were interested to test whether NPR3 and NPR4 also bind to the PR1 promoter in a manner similar to NPR1, in presence and absence of SA. To do so, similar to our previous approach, modified ChIP experiments were conducted to analyse the binding of both NPR3 and NPR4 to the -701 and -870 fragments of the PR1 promoter before and after addition of 1μM SA (Figure 16A). For these experiments, NPR1 was employed as a positive control measure.

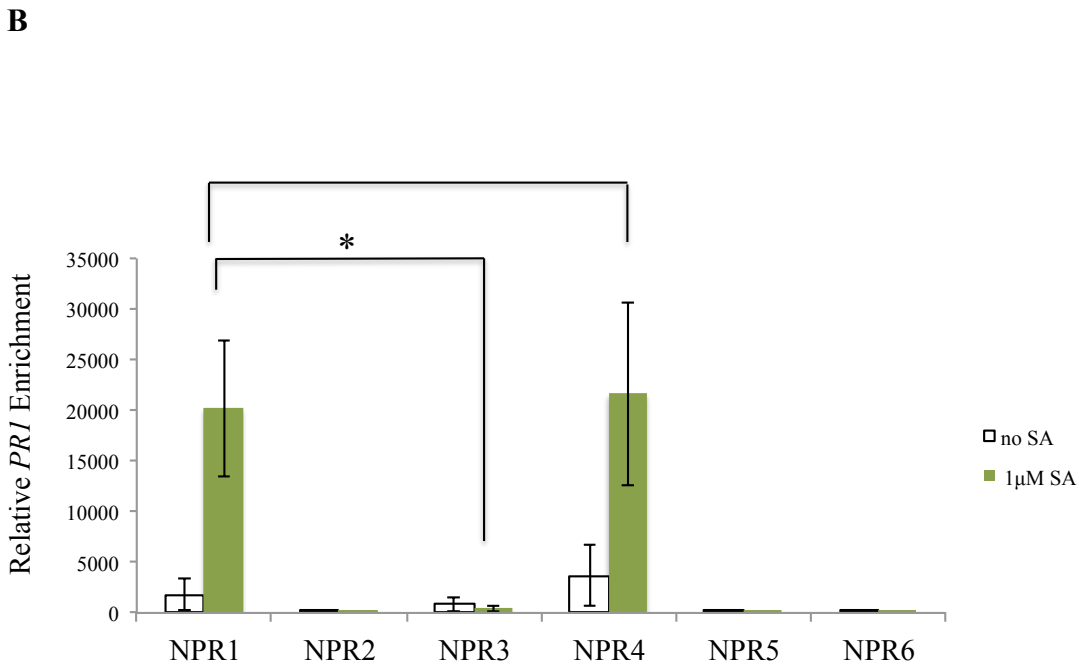
Similar to NPR1, our results indicated that NPR4 binds to the -701 fragment only in the presence of SA, and to the -870 fragment only in the absence of SA (Figure 16B and C). Conversely, NPR3 was shown to evade binding to the -701 fragment, neither in the presence nor in the absence of SA, although binding of NPR3 to the -870 fragment was shown to occur in absence of SA (Figures 16B and C). Student two-tailed t-test results showed that statistically significant differences were not observed between NPR1 and NPR4 relative *PR1* enrichment values obtained in the presence of SA (p-value is 0.8500; p-value < 0.05). However the relative *PR1* enrichment value obtained for NPR3 was significantly reduced (Figure 16B) (p-value is 0.0291; p-value < 0.05) compared to both NPR1 and NPR4. Additionally, no

significant differences were observed between the amounts of relative *PR1* enrichment obtained for NPR1, NPR3, and NPR4 for the -870 fragment in the absence of SA (Figure 16 C) (Student two-tailed t-test, p-values are 0.1021, and 0.1622 respectively; p-value < 0.05).

Chiefly, these results indicate that both NPR3 and NPR4 act as positive regulators at different stages of SAR establishment, similar to NPR1.

Similar to our previous approach, the binding of NPR2, NPR5, and NPR6 to the -701 and -870 fragments of the PR1 promoter, before and after addition of 1μM SA, were also analyzed through the employment of the modified ChIPs procedure (Figure 16A). Unsurprisingly, both NPR5 and NPR6 failed to display a binding preference towards either the -701 or the -870 fragments of the PR1 promoter, consistent with their established function in plant developmental processes (Hepworth et al., 2005; Norberg et al., 2005) and lack of evidence supporting plant defense-related functions (Figure 16B and C). NPR2, on the other hand, bound to the -870 fragment in the absence of SA, but not to the -701 fragment (Figure 16B and C). These results are consistent with the fact that NPR2 lacks the amino acids cysteine 521 and 529, which have been shown to play critical roles in the capture of SA (Figure S6).


A
-1023
WT PR1 promoter -1 to -1023
-1
-1023
-870
PR1 promoter -1 to-870 (Δ-870-1023)
-1
-1023
-701
PR1 promoter -1 to -701 (Δ-701-1023)
-1
NPR1 (1-593)
NPR2 (1-600)
NPR3 (1- 586)
NPR4 (1-574)
NPR5 (1-491)
BOP2
NPR6 (1-467)
BOP1


B
*
Relative PR1 Enrichment
35000
30000
25000
20000
15000
10000
5000
0
NPR1
NPR2
NPR3
NPR4
NPR5
NPR6
no SA
1μM SA

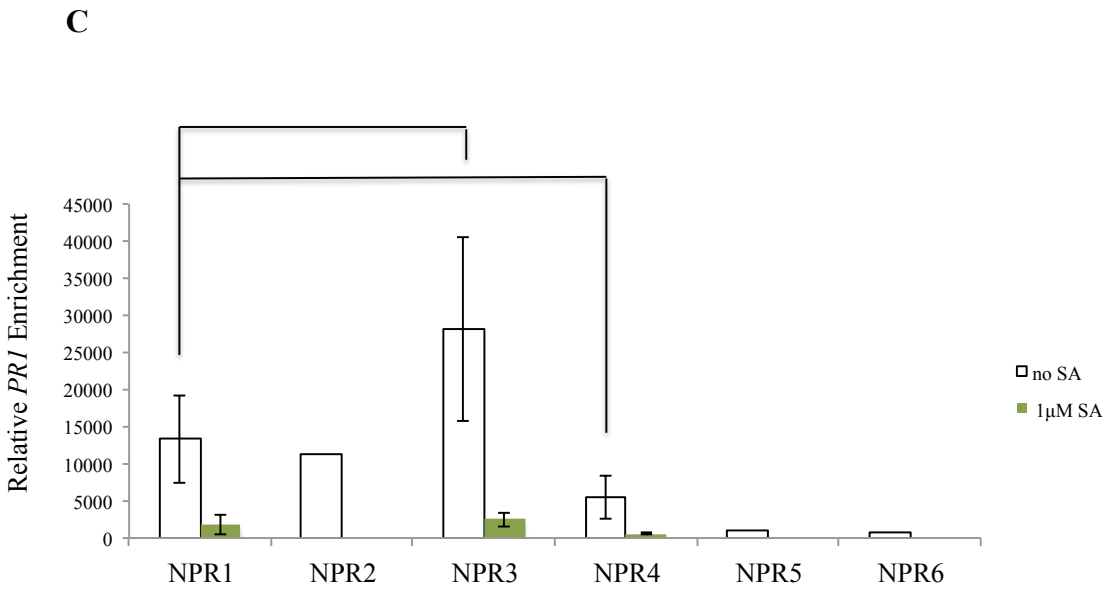


Figure 16. Binding of NPR2, NPR3, and NPR4 to the PR1 promoter

A. Schematic representation of the PR1 promoter deletions constructs and *Arabidopsis* NPR1-like proteins used for modified ChIP experiments. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). White box indicates the -1 to -870bp fragment of the PR1 promoter, which sequences from -870bp to -1023bp deleted (Grey colour box). Green box represents the -1 to -701bp fragment of the PR1 promoter that sequences from -701bp to -1023bp deleted (Grey colour box). The arrow indicates the start of transcription. White boxes show all six members of the *Arabidopsis* NPR protein family, NPR2, NPR3, NPR4, NPR5 (BOP2), and NPR6 (BOP1). Regions with homology to known functional domain are shown in different color boxes. Hatched boxes represent BTB/POZ domains, grey boxes show ankyrin repeat regions, and Black boxes indicate nuclear localization sequence.

B. Histogram illustrating modified ChIP experiments performed on *Arabidopsis* NPR proteins and the -701 fragment of the PR1 promoter before and after addition of 1μM SA, confirming that similar to NPR1, NPR4 binds to the -701 fragment in the presence of SA. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

C. Histogram illustrating modified ChIP experiments performed on *Arabidopsis* NPR proteins and the -870 fragment of the PR1 promoter before and after addition of 1μM SA, confirming that NPR2, NPR3, and NPR4 bind to the -870 fragment in the absence of SA in a manner similar to NPR1. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

4.12 Binding of NPR1 protein to the PR1 promoter for *Zea mays* (corn) and *Oryza sativa* (Rice)

In the last decades, several studies reported that SA-dependent, NPR1-mediated defense responses are not only found in *Arabidopsis,* but also conserved among other plant species. Therefore, similar to our approach for *Arabidopsis* NPR1, the binding of corn and rice NPR1 proteins were tested for both the -701 and -870 fragments of the PR1 promoter, before and after addition of 1μM SA (Figure 17A).

Our findings revealed that the bindings of both corn and rice NPR1 proteins to the -701 fragment after addition of SA yielded the largest relative *PR1* enrichment values (Figure 17B). Student two-tailed t-tests showed no significant differences between relative *PR1* enrichment values found for corn and rice before and after SA addition (p-values are 0.2340, 0.7568 respectively p-value < 0.05) (Figure 17B). In addition, corn and rice NPR1 did not bind to the -870 fragment of the PR1 promoter (Figure 17B).

Considering that in the presence of SA, both corn and rice NPR1 proteins showed similar binding profiles as *Arabidopsis* NPR1 to the -701 fragment, we were interested to determine whether the recruitment of corn and rice NPR1 proteins to the PR1 promoter is also metal-dependent, similar to *Arabidopsis* NPR1 (Figure 11C). To do so, similar to our previous approach (see section 4.8), EDTA was employed as a metal chelating agent and used to incubate NPR1 (corn and rice individually) and the -701-fragement of the PR1 promoter (10mM EDTA) prior to cross-linking. Surprisingly, the relative *PR1* enrichment processes for both corn and rice NPR1 proteins were completely abolished after EDTA addition (Figure 17B), confirming that the recruitment of both corn and rice NPR1 proteins to the PR1 promoter is also metal-dependent.

Next, the bindings of corn, rice, and *Arabidopsis* NPR1 proteins to the -701 fragment (Figure S8A) were tested before and after addition of 1μM SA and 1μM Isotianil (Figure S8B). Similar to SA, Isotianil has been shown to be involved in SAR establishment in rice plants (Ogawa et al., 2011). Consistent with previous results, (Figure 17B) our findings showed that corn, rice, and *Arabidopsis* NPR1 proteins bound to the - 701 fragment in the presence of SA (Figure S8C). However, no *PR1* enrichment was observed after Isotianil addition, perhaps due to solubility issues.

Although corn, rice, and *Arabidopsis* NPR1 proteins only share a low percentage of sequence identity (Table S2 and Figure S7), and are not found close to one another in the phylogenetic tree (Figure S5B), surprisingly, our results suggest that the recruitment of NPR1 to the PR1 promoter dependent on the presence of SA and metal is conserved in corn, rice, and *Arabidopsis*.

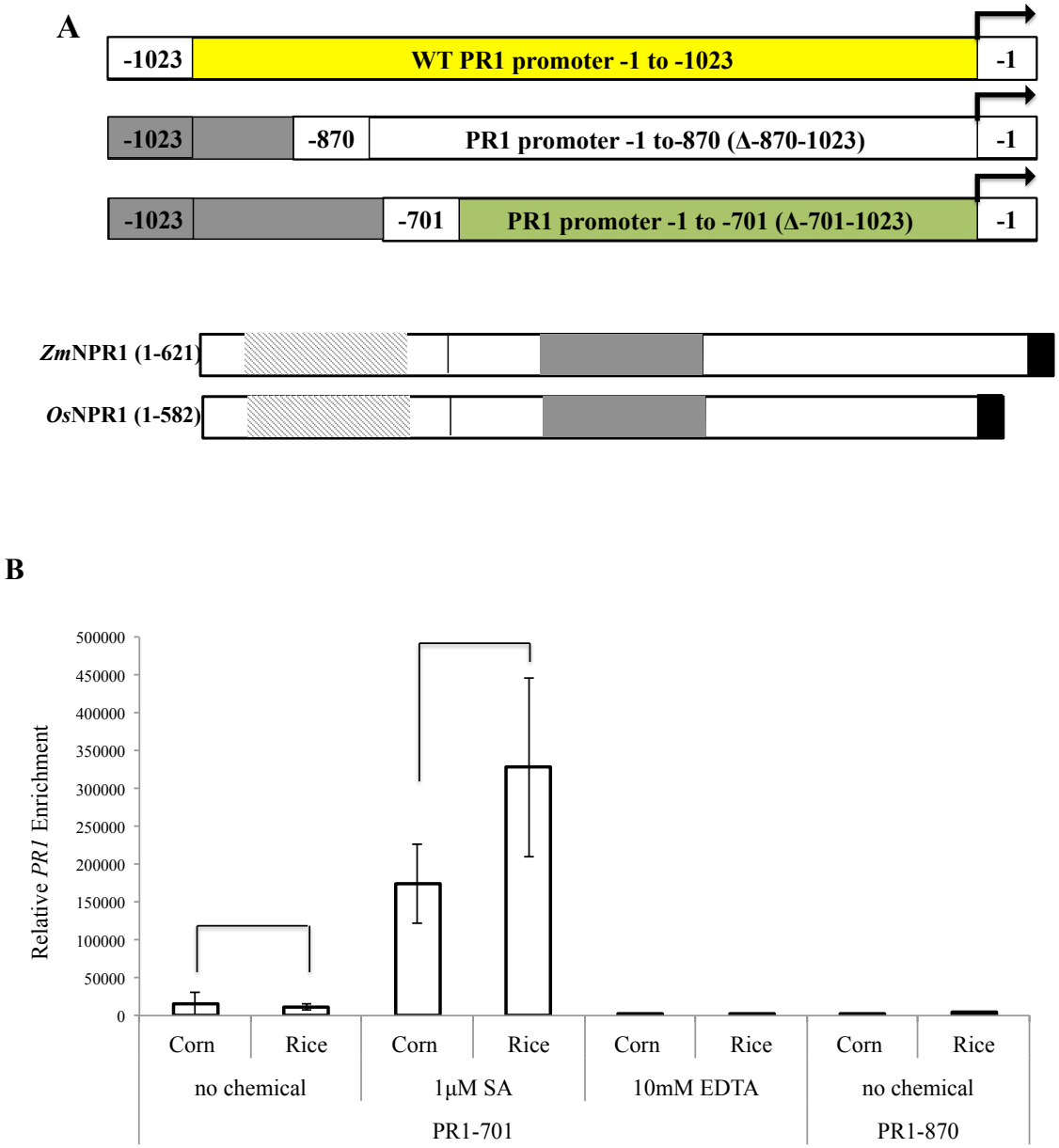


Figure 17. Conservation of metal and SA-dependent recruitment of NPR1 to the -701 fragment of the PR1 promoter for rice and corn

A. Schematic representation of the PR1 promoter deletions constructs and NPR1 proteins from maize and rice used for modified ChIP experiments. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). White box indicates the -1 to -870bp fragment of the PR1 promoter, which sequences from -870bp to -1023bp deleted (Grey colour box). Green box represents the -1 to -701bp fragment of the PR1 promoter that sequences from -701bp to -1023bp deleted (Grey colour box). The arrow indicates the start of transcription. White boxes show *Zea mays* NPR1 protein (*Zm*NPR1) (amino acids 1-621), and *Oryza sativa* NPR1 protein (*Os*NPR1) (amino acids 1-582). Regions with homology to known functional domains are shown in a different color box. Hatched boxes represent BTB/POZ domains, grey boxes show ankyrin repeat regions, and Black boxes indicate nuclear localization sequence.

B. Histogram illustrating modified ChIP experiments performed on rice and corn NPR1 proteins with the -701 and -870 fragments of the PR1 promoter, in the presence and absence of 1μM SA and 10mM EDTA, confirming that both rice and corn NPR1 proteins bind to the -701 fragment in the presence of SA and metal, although neither were shown to bind to the -870 fragment. ChIPs were performed with anti-Strep antibody. Relative *PR1* enrichment values were analysed by qPCR. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates. The asterisk (*) indicates the statistical significance, assessed using a two-tailed Student's t-test (p values <0.05).

CHAPTER 5 – DISCUSSION

In recent years, the roles of NPR1 and the SA-signaling network in plant defense mechanisms have been the focus of many groups of research (Fu et al., 2012; Manohar et al., 2015; Wu et al., 2012; Zhang et al., 2006). These studies have been able to demonstrate that NPR1 plays a key role in plant immunity by regulating both local and systemic acquired resistance (SAR). As such, plants that lack the functional NPR1 protein have been shown to be unable to activate SAR or express the SAR marker gene, *pathogenesis related 1 gene* (*PR1*); consequently, they cannot survive attacks stemming from biotrophic pathogens.

In order to further elucidate whether NPR1 is capable of forming a complex with TGA2 on the PR1 promoter to regulate *PR1* expression, Rochon et al. (2006) performed in vivo ChIP experiments on wild type, *npr1-3* (Cao et al., 1997), and *tga2/5/6* mutant *Arabidopsis* plants prior to and after treatment with SA. In this work, NPR1 was shown to interact with the PR1 promoter both prior to and after addition of SA, and independently of TGA2/5/6 (Figure S1). However, due to insufficient experimental evidence, these findings failed to confirm that NPR1 has a DNA binding domain that allows its direct recruitment to the PR1 promoter.

5.1 Validation of modified in vitro Chromatin Immunoprecipitation (ChIP) method

Currently, a host of experimental methods are available that allow for the analysis of protein–DNA interactions. One of these methods is chromatin immunoprecipitation (ChIP), which enables the in vivo capture of protein–DNA interactions (Collas and Dahl, 2008; Das et al., 2004).

As the traditional in vivo ChIP method is not applicable to study the binding of NPR1 to the PR1 promoter, a modified in vitro ChIP method had to be implemented and validated. The resulting method presents a host of advantages in regards to its application towards the characterization of the NPR1-PR1 promoter binding. For instance, the method enabled the direct analysis of the PR1 promoter sequence, as sequences were cloned into a plasmid. In addition, the modified ChIPs method also allowed for direct analysis of the NPR1 protein, which is expressed as an *E. coli* recombinant protein. As well, the method did not require the use of sonication to breakdown the chromatin, as the PR1 promoter sequence was easily accessible in a plasmid for protein interaction. In traditional ChIP methodology, sonication efficiency is a critical step in the successful implementation of the method; if the chromatin is overly sheared, or does not completely break, false positive or negative results may occur (Haring et al., 2007; Mukhopadhyay et al., 2008). For the purposes of our study, an *E. coli*-expressed streptavidin-tagged protein was used without any purification as protein input in our modified ChIP protocol. This presented two clear advantages; by avoiding the purification step, we increased the amount of available recombinant protein to bind to the DNA. This provided a distinct benefit to the method, as the amount of expressed protein may sometimes be decreased during the purification step, particularly for low-expressed proteins tagged to a specific amino

acid peptide. Secondly, the absence of a purification step provided a competition state (Lickwar et al., 2012) between our target protein and *E. coli* proteins for binding to the PR1 promoter, which further confirmed the specificity and robustness of the interaction.

Conversely, the classical in vivo ChIP technique is burdened with a few disadvantages; for example, as generally is the case for in vivo methodology, it suffers from complications in variability (Mukhopadhyay et al., 2008; Park, 2009). Other classical protein-DNA interaction techniques, such as Electrophoretic Mobility Shift Assay (EMSA), also suffer from drawbacks associated to false negatives (Carey et al., 2012; Hellman and Fried, 2007); this is generally observed due to the complicated nature of the NPR1- PR1 promoter interaction.

The first major challenge that had to be overcome in the development of the modified ChIP technique was the identification of a protein candidate as a positive control whose recruitment to the PR1 promoter had already been well-established in the literature. The PR1 promoter in *Arabidopsis* has 2 cis-acting elements with TGACG binding motifs, linker-scanning 7 (LS7) and linker-scanning 5 (LS5). TGA2 transcription factor binds both elements with equal affinity in vitro (Boyle et al., 2009; Johnson et al., 2003; Pape et al., 2010). Thus, for development and validation of the in vitro ChIP method, the full-length TGA2 protein was chosen as our positive control candidate, while a shorter version of TGA2 (VFI-TGA2), which lacks the domain necessary for DNA binding, was chosen as our negative control candidate (Figure 1A). In addition, a multi-step method optimization was undertaken, including optimization of parameters such as concentrations of input DNA, protein samples, antibody, reference DNA, as well as protein-DNA crosslinking.

In order to optimize the crosslinking of the TGA2 protein with the PR1

promoter, various incubation times and different concentrations of formaldehyde were tested. Our results revealed TGA2 was recruited to the PR1 promoter when a concentration of 0.01% formaldehyde was used, and an incubation time of 5 min was set for crosslinking (Figure 1B). The TGA2 *E. coli* expressed proteins presented in a crude extract could easily and directly access the sequence of the PR1 promoter presented in a plasmid. As such, increases in fixative time and concentrations could raise non-specific false positive binding results. Because longer crosslinking time can favors protein-protein interactions and also the epitope recognized by antibody sometimes can be masked that eventually result in loss of immuno-precipitation signals (Jain et al., 2015; Teytelman et al., 2013). To illustrate, as can be seen from Figure 1B, when higher concentrations of formaldehyde were used (either 1% or 0.1%) no differences were observed between positive and negative controls (VFI-TGA2 and TGA2 respectively). However, at a 0.01% formaldehyde concentration level, a distinct difference could be observed between the *PR1* enrichment levels of VFI-TGA2 and TGA2 proteins.

The quality and specificity of the selected antibodies play a very important role in the immunoprecipitation of the cross-linked DNA-protein complex (Haring et al., 2007; Mukhopadhyay et al., 2008). Compared to polyclonal antibodies, monoclonal antibodies have much higher specificity and consistency; contrastingly, polyclonal antibodies boast the advantage of recognizing several epitopes on a given target, and hence, they increase signal levels of low-abundance templates (Perez-Burgos et al., 2004). With these considerations in mind, polyclonal anti-Strep and anti-His antibodies were selected as immunoprecipitators for modified in vitro ChIP analyses. The specificity and TGA2 and/or NPR1-recognition ability of the anti-Strep and anti-His antibodies were verified with the use of the western blotting procedure.

Furthermore, to confirm the suitability of the anti-Strep antibody as the immunoprecipitator of choice, modified ChIP experiments were also performed with the use of anti-His antibody as an immunoprecipitator of the strep-tag protein bound DNA complex. The obtained results further confirmed that the anti-Strep antibody specifically immunoprecipitates the protein bound DNA complex, as no *PR1* enrichment was detected upon use of anti-His antibody within the same set of samples.

After confirmation of the specificity of the anti-Strep antibody, a series of antibody dilutions were prepared and tested for the detection of *PR1* enrichment in modified ChIP experiments. The obtained results revealed that a 1:200 dilution (0.00125mg/ml) of antibody in IP buffer offered an optimum concentration of antibody for the immunoprecipitation of the protein-bound DNA complex (Figure 1C). However, at higher antibody concentrations (1:50 dilution), non-specific binding was observed for VFI-TGA2 compared with TGA2 (Figure 1C).

5.2 Determination of multiple NPR1 binding sites on the PR1 promoter by validated in vitro ChIP method

Following the conclusion of its optimization (see sections 4.1 and 5.1), the modified in vitro ChIP method was applied towards the binding of NPR1 to the PR1 promoter. Relative *PR1* enrichment amounts similar to those obtained for TGA2 protein were obtained for NPR1 (Figure 1D), confirming that NPR1 was recruited to the PR1 promoter directly. This result was in agreement with previous results obtained by Rochon et al. (2006), whose in vivo work demonstrated that NPR1 is recruited to the PR1 promoter.

In order to identify the location where NPR1 is recruited on the PR1 promoter, a series of PR1 promoter deletion constructs were used to test the binding of NPR1. Our obtained results showed that NPR1 binds around the -870 fragment region (Figure 2A) (-1 to -870bp) of the PR1 promoter (Figure 2B). Conversely, no binding activity was observed around the -621 fragment region (-1 to -621bp) (Figure 2A and B). Both the absence of binding activity on the -621 fragment and the presence of binding on the -870 fragment suggested that NPR1 might have multiple binding sites on the PR1 promoter. Our results support the findings obtained by Wu et al. (2012), who demonstrated that different domains of NPR1 are involved in the SA-mediated signalling switch of SAR and *PR1* gene activation.

Similar to the results obtained in this work, the existence of multiple binding sites on promoters has been reported for different DNA-binding proteins. Examples include related to ABI3/VP1 (RAV1) and (RAV2) transcription factors, associated with ethylene sensing and floral induction of *Arabidopsis* (Kagaya et al., 1999; Matias-Hernandez et al., 2014), and telomeric repeat binding factor 1 (TRF1), a

duplex telomeric DNA binding factor found in mammalian chromosomes (Bianchi et al., 1999; Chong et al., 1995).

The results obtained for this series of experiments (see section 4.1 and 4.2) confirmed that NPR1 binds to the PR1 promoter directly, and that the modified in vitro ChIP method is applicable to study complicated protein-DNA interactions. As a next step, the role of SA in the binding of NPR1 to the PR1 promoter was further explored through a series of modified ChIP experiments.

5.3 Binding of NPR1 to the -701 and -870 fragments of the PR1 promoter in the presence and absence of SA

Two independent groups, Wu et al. (2012) and Manohar et al. (2015), have shown that NPR1 is a SA-specific receptor. As such, we elected to test the effect of SA on the binding of NPR1 to the PR1 promoter. Our results demonstrated that SA affects NPR1 binding sites in a contrasting manner (Figure 3C). In the -870 fragment of the PR1 promoter, the presence of SA inhibited the binding of NPR1 to the PR1 promoter, while for the -701 fragment, SA promoted NPR1-PR1 promoter interactions (Figure 3C). Our findings were consistent with previous data (Figure 2B), and further supported the indication that NPR1 contains multiple binding sites on the PR1 promoter.

Next, we applied a molecular approach to identify specific NPR1 binding recognition sites through the dissection of the PR1 promoter (Ito et al., 2005; Lehmeyer et al., 2015). Although the deletion analysis approach allowed for recognition of important NPR1 recruitment regions (Figure 2), further analysis of a series of PR1 promoter deletion constructs (Figure 4A and 5A) allowed for higher resolution identification of short and defined sequences of the PR1 promoter. Our findings revealed that in the presence of SA, NPR1 is specifically recruited at the -636 to -646 base pair sequence of the PR1 promoter (Figure 4); conversely, in the absence of SA, NPR1 binding was shown to occur at the -790 to -833 fragment of the PR1 promoter (Figure 5). Our results were also in agreement with the findings presented by Wu et al. (2012). They performed 3C *in vivo* experiments to study NPR1 conformations. Only in absence of SA NPR1 oligomer was formed on PR1 promoter with enrichment around -734 to 833bp fragment.

A similar molecular promoter dissection approach has also been applied to

study promoters of *PR* genes in different plant species. For example, in work conducted by Li et al. (2005), the -362 to +106 fragment of the PR1 promoter of barley (*Hordeum vulgare*) was demonstrated to be critical for SA-responsive signaling. Similarly, the -273 to -226 and -743 to -707 fragments of the PR1 promoter region of parsley (*Petroselinum crispum*) were demonstrated to be essential for SA-responsive *PR1* expression (Rushton et al., 1996). In other work, similar regions on the -687 to -637 fragment of the rice PR10a promoter have also been indicated to be critical for SA induction of *PR10a* expression (Hwang et al., 2007).

The obtained findings of Rushton et al. (1996) and Hwang et al. (2007) in regards to parsley and rice PR1 promoters were further validated through the results obtained in the current work, as the same regions, around the -636 to -645 base pair, were shown to be important for *Arabidopsis* NPR1 binding in the presence of SA (Figure 4). Although the exact location of the identified regions differ slightly between our findings and those of Rushton and coworkers (-636 to -645 vs. -743 to -707, respectively), and Hwang and coworkers (-636 to -645 vs. -687 to -637, respectively) a close consistency between them can still be observed.

Our results were also in agreement with the findings presented by Lebel et al. (1998). Their results were further validated through our research by demonstrating that NPR1 binds to the -636 to -646 fragment region in the presence of SA and to the -790 to -833 fragment in its absence. In their work, the -636 to -645 and -815 fragments on the PR1 promoter were shown to be critical for *PR1* expression in both the presence and absence SA.

To reiterate, the -636 to -645 and -790 to -833 base pair regions were confirmed through our research as the sites on the PR1 promoter where the binding of NPR1 occurs, dependent on the presence or absence of SA. As a next step, we sought

to determine whether the active analogs of SA, namely 4-OH-BA and INA, affect the binding of NPR1 to the PR1 promoter.

5.4 Negative regulation of NPR1 binding to PR1 promoter by INA and 4-OH-BA

As previously discussed in the literature review, SAR-mediated plant immunity has been shown to be induced when a local infection occurs. The onset of SAR requires the accumulation of SA and the corresponding expression of *PR* genes. Exogenous application of SA activates *PR* gene expression and resistance in plants without pathogen inoculation. A wide range of synthetic chemicals have been shown to be able to mimic SA as a natural inducer of plant defence responses, such as 2,6-Dichloroisonicotinic acid (INA) and Benzo-(1,2,3)-thiadiazole-7-carbothioic S-methyl ester (BTH) (see section 2.1.5.3) (Friedrich et al., 1996; Kessman et al., 1994; Métraux et al., 1991; White, 1979). In tobacco plants, both SA and INA have been shown to induce *PR* genes. However, our results revealed that INA is unable to stimulate NPR1 binding to the PR1 promoter in a similar manner as SA, suggesting that INA might not act through the same pathway as SA. Two halogenated and aldehyde forms of INA were also shown to activate *PR1* expression in tobacco; however, other INA analogues were not active at a similar high concentration (1mM). Similarly, the SA derivatives 3-hydroxybenzoic acid (3-OH-BA) and 4-hydroxybenzoic acid (4-OH-BA) did not activate *PR1* expression at similar concentrations (1mM)(Conrath et al., 1995).

Wu and co-workers (2012) demonstrated that the SA derivatives 3-OH-BA (Conrath et al., 1995), 4-OH-BA (Bi et al., 1995), and INA were poor inducers of *PR1* expression in *Arabidopsis*. Particularly, they reported that INA was 42 times less effective than an identical concentration of SA (300μM). Furthermore, through the application of equilibrium dialysis competitive binding assays, they demonstrated that the SA derivatives 3-OH-BA, 4-OH-BA, and INA do not bind NPR1 with the same affinity as SA, consistent with their capacity as poor inducers of *PR1* expression.

The discovery that 4-OH-BA and INA failed to interact with NPR1 during equilibrium dialysis competitive binding assays prompted us to investigate their effect on NPR1 binding to PR1. In contrast to SA, INA and 4-OH-BA, administered at two different concentrations (0.1 and 1μM), completely abolished NPR1 binding to the -701 and -870 regions of the PR1 promoter (Figure 3D). These results are consistent with previous results obtained by Wu and co-workers (2012), and confirm the specificity of the effect of SA on direct NPR1 binding to the PR1 promoter. In addition, our results demonstrated that both 4-OH-BA and INA negatively regulate NPR1 interaction with the PR1 promoter.

Our findings in regard to the abolishment of NPR1 binding interaction upon addition of INA serve to clear some of the ambiguity arising from contradictory past results reported by different groups on *PR1* expression by INA (Cao et al., 1997; Wu et al., 2012; Zhang et al., 2003). Further, our findings confirm the conclusions made by Wu et al. (2012) and Conrath et al. (1995), who have suggested that INA activates *PR1* through a mechanism different from that of SA.

As explained by Wu et al. (2012), the structural differences between SA, INA, and 4-OH-BA could explain why INA and 4-OH-BA behaved differently from SA in terms of binding to NPR1, and also why they negatively affected NPR1 interaction with the PR1 promoter. Since BTH showed a similar, slightly higher affinity for NPR1 binding in equilibrium dialysis binding assays as compared to SA, Wu et al., (2012) described that a hydroxyl group in ortho position to a free carboxylic acid on the aromatic ring would be the essential elements required for NPR1 binding, and perhaps also for affecting its interaction with the PR1 promoter. Indeed, this hypothesis is further supported by the fact that INA lacks the needed hydroxyl group, which according to the findings described by the authors, would be a critical factor for

the binding of these chemicals to NPR1, which would in turn affect the binding of NPR1 to the PR1 promoter.

5.5 NPR1 only binds circular double stranded form of DNA

After successfully determining the specific binding sites of NPR1 on the PR1 promoter in both the presence and absence of SA (Figure 4 and 5), and disclosing the locations of the NPR1 DNA binding domains (Figure 8 and 11), we set out to determine the DNA binding properties of NPR1. To do so, simple restriction digestion experiments (Figure S9) on both the -701 and -870 fragments of the PR1 promoter (Figure 12A) were conducted in the presence and absence of SA and with the use of two different restriction enzymes, *Bgl*I and SnaBI. NPR1 was shown to only be recruited to the PR1 promoter when DNA was undigested and in circular form, both in the presence and absence of SA (Figure 12B and C). In this work, the binding preference of NPR1 was also addressed in regard to its preference towards double-stranded versus single-stranded DNA. For this determination, sample DNA was submitted to denaturation by boiling, followed by analysis of relative *PR1* enrichment values. Our results revealed that NPR1 recruitment was significantly reduced or even abolished for both the -701 and -870 fragments of the PR1 promoter (Figure 12B).

This is not the first example of a plant DNA-binding protein that only binds to circular double-stranded DNA. This characteristic of NPR1 certainly helps to explain why NPR1 had not been identified as a DNA-binding protein to date. Many researchers, by default, have used the linear DNA form in common DNA-protein interaction techniques (Tsai et al., 2012), including Electrophoretic mobility shift assay (EMSA). However, as demonstrated in the current work, the recruitment of NPR1 to the PR1 promoter requires circular DNA, while linear DNA precludes NPR1 from binding to the PR1 promoter.

Few other reports have also demonstrated the existence of a specific binding preference of certain proteins for circular-form DNA. For instance, Kant et al. (2005) demonstrated that the rice DNA Meiotic Recombinase 1 protein (*Os*DMC1), exhibits a binding preference towards circular single-stranded DNA (ssDNA) and circular double-stranded DNA (dsDNA). In other work, conducted by Rojas and co-workers (1998), bean dwarf mosaic geminivirus movement proteins were shown to exhibit a binding preference for open circular forms of DNA. As emphasized by the authors, these binding preferences have been shown to be critical for bending of the DNA molecule, which is followed by folding and/ or twisting, resulting in the formation of a stabilized protein-DNA complex. Liu et al. (1997) have also demonstrated that the Maize streak virus (MSV) coat protein (CP), required for virus movement within the plant, binds to single-stranded and double-stranded viral and plasmid DNA in a non-specific sequence manner.

The Ascorbate Oxidase gene-binding protein (AOBP) has been recognized as a protein that binds to the unique sequence AGTA repeat found in the silencer region of the *Arabidopsis* aldehyde oxidase (AAO) gene promoter (Kisu et al., 1998). Similar to our findings, which have shown that NPR1 binds exclusively to double-stranded fragments of the PR1 promoter, Umemora et al. (2004) have shown that the AOBP with Dof (DNA-binding with one finger) domain specifically binds to a double-stranded AGTA repeat, but not to the single-stranded sense or anti-sense AGTA repeat.

Double stranded DNA (dsDNA) is composed of a uniform structure along with a highly negatively charged sugar-phosphate backbone, and stacked base pairs with edges containing major and minor grooves. Each DNA sequence contains a chemical signature described by the pattern of the base pairs exposed in the DNA

grooves (Luscombe et al., 2001). Proteins recognize their interacting DNA partners either through the unique chemical signatures of the DNA bases (base readout) or by a sequence-dependent DNA shape (shape readout) (Rohs et al., 2010).

In the case of NPR1 binding to the PR1 promoter, we were interested to find out what recognition mechanisms or factors play a role in the interaction. Careful study of the PR1 promoter sequence led to the discovery of an Adenine/Thymine (A/T)-rich fragment around the -700 to -800 base pairs. Interestingly, these sites were found to be in close proximity to the sites where NPR1 binds to the PR1 promoter in the presence (636 to 646 base pair sequence) and absence (790- to 833 base pair sequence) of SA. In the literature, increased expression of the *β-glucuronidase* (GUS) reporter gene has been shown to occur in transgenic tobacco when an A/T-rich element from the promoter of the French bean *P-phaseolin* gene was fused to the cauliflower mosaic virus 35s promoter (Bustos et al., 1989). Additional evidence supporting a role for A/T rich sequences in transcriptional activation was also demonstrated for the soybean heat shock promoter (Czarnecka et al., 1992), and in peas, within the small sub-unit of the *ribulose bisphosphate carboxylase (rbcS-3A)* gene promoter (Lam et al., 1990). The abovementioned findings from the literature, together with our results, suggest that the A/T-rich fragment around the SA-induced NPR1-PR1 promoter binding site, similar to the A/T fragment found in pea and soybean promoters, may be involved with NPR1 transcriptional activity and *PR1* activation in the presence of SA. Conversely to these reports of A/T-rich elements associated with transcriptional activity, Castresana et al., (1988) and Datta and Cashmore, (1989) identified an A/T rich sequence that acts as a negative regulatory element in the *Nicotiana plumbaginifolia* chlorophyll alb binding protein (Cab-E) gene promoter. Clearly, the evidence that A/T-rich fragments play an additional role

as negative regulatory elements helps to further understand the role of the found A/T-rich fragment in the PR1 promoter close to where NPR1 binds in the absence of SA.

Given that NPR1 only binds circular and double-stranded forms of DNA and that this interaction only occurs in the presence of an A/T-rich fragment, we conducted an investigation into the formation of DNA secondary structures, which could help to stabilize the NPR1-PR1 promoter interaction. With the use of Mfold DNA secondary structure prediction software (Zuker, 2003), stem-loop (hairpin) structures (Figure S3) were identified. Stem-loop secondary structures are inverted repeat sequences known to form in negatively supercoiled DNA (Wells, 1988; Sinden, 1994). They are widespread in the genomes of both eukaryotes and prokaryotes, occurring more often than expected among random sequences (kato et al., 2003). The formation of such stem-loops around the sites where NPR1 binds to the PR1 promoter could possibly enhance or stabilize the NPR1-PR1 promoter complex. Nevertheless, a classical *S1* endonuclease enzyme mapping assay did not confirm the formation of secondary structures (data not shown).

In addition to α helices and β sheets of proteins, which provide a very solid scaffold for side- and main-chain interactions with DNA, loops are also known to play a role as recognition elements between DNA and proteins. Well-established examples of DNA-binding proteins that recognize their DNA partners through loops have been demonstrated for the superfamily of DNA-binding proteins that contains an immunoglobulin-like fold protein (Garvie and Wolberger, 2001). Within our results, a formation of multi-loop structures on the PR1 promoter was also observed in the sites where NPR1 binds in the presence and absence of SA. Particularly, upon exploration of the fragment region around the -636 to -646 base pair sequence, where NPR1 has been shown to interact with the PR1 promoter in the presence of SA, a formation of

multi-loops was identified and hypothesized to contribute to the stabilization of the NPR1-PR1 promoter complex (Figure S3A). Interestingly, upon deletion of sequence base pairs -636 to -646 (LS7 site mutation), the formation of one of these multi-loops was observed to be disrupted, which resulted in complete abolishment of NPR1 binding to the PR1 promoter (Figure S3A and Figure 4). Similarly, the disruption of in-loop formations around the -790 fragment resulted in complete abolishment of NPR1 recruitment to the PR1 promoter (Figure S3B and Figure 5). These results further confirmed the importance of DNA secondary structure formations, including stem and multi-loops for the stabilization of protein-DNA interactions.

5.6 Localization of NPR1 DNA binding domains in the N and C-terminus regions

Previous to the current research, the identity of NPR1 as a DNA-binding protein was unknown. Indeed, numerous publications indicated that NPR1 does not have a DNA binding domain (Kuai et al., 2015; Rochon et al., 2006). Its role had been relegated to the stimulation of DNA binding activity of interacting TGA factors to SA-response elements (Després et al., 2000; Fan and Dong, 2002; Johnson et al., 2003). In spite of these reports, NPR1 has been demonstrated to be recruited to PR1 (see sections 4.1, 4.2, 4.4 and 4.5). In addition, two distinct DNA-binding domains within the N (amino acid 110-190) and C-terminus (amino acid 513-535) regions of NPR1 (Figure 11 and 8 respectively) were demonstrated to exist in the current research. The results obtained from this series of experiments (see sections 4.6, 4.7, and 4.8) further support previous data (see sections 4.2, 4.4 and 4.5) that suggested that NPR1 has two distinct binding profiles, in the presence and absence of SA, at two different sites on the PR1 promoter. As such, two distinct DNA binding domains exist for the binding of NPR1 to individual recognition sites on the PR1 promoter.

The existence of a DNA binding protein with two distinct DNA binding domains is not an unusual occurrence. Similar profiles have been reported for a variety of DNA binding proteins, including *Arabidopsis* related to ABI3/VP1 (RAV1) (Hu et al., 2004), zinc transport of *Arabidopsis* 10 (Zat10) (Sakamoto et al., 2004) and Zat12 (Davletova et al., 2005; Iida et al., 2000; Rizhsky et al., 2004; Vogel et al., 2005), *Arabidopsis* zinc finger proteins 1 (AZF1), 2 (AZF2), and 3 (AZF3) (Sakamoto et al., 2002), wheat leaf rust resistance 10 (Lr10) (Feuillet et al., 1997), barley Ruptured pollen grain 1 (Rpg1) (Brueggeman et al., 2002; Mirlohi et al., 2008), and rice blast pathogen (Pib) (Wang et al., 1999). Consistent with our results on NPR1, the identified DNA binding domains for *Arabidopsis* CONSTANS (CO) B-

box protein (Khanna et al., 2009; Putterill et al., 1995) and stress-associated protein (SAP) in rice and *Arabidopsis* have been shown to be localized in both the C and N terminals (Mukhopadhyay et al. 2004; Vij et al. 2006).

An exploration into the C-terminus region of NPR1 was first conducted so as to identify the DNA binding domain of NPR1 (Figure 6 and 7). In this group of experiments, the DNA binding domain of NPR1 in the presence of SA was first narrowed down to two different sites on the C-terminus region. These regions were localized within the amino acid sequences 374 to 593 (Figure 6) and 463 to 570 (Figure 7). Despite the observed inconsistencies in identified regions in the C-terminus as a potential DNA binding domain, amino acid sequence composition analysis revealed the presence of charged amino acids, including histidine (H), arginine (R), lysine (L), and serine (S), on the C-terminus region. These results support the theory that particular amino acids-base pairs are favored in direct recognition of DNA base sequences by proteins (Luscombe et al., 2001; Rohs et al., 2010). For example, the amino acids arginine, lysine, serine, and histidine have been shown to preferentially interact with guanine, while the amino acids asparagine and glutamine preferentially interact with adenine (Luscombe et al., 2001; Suzuki, 1994).

It is possible that the observed inconsistencies were due to improper folding and conformational instability of the proteins. Since the NPR1 C-terminus deletions constructs were made either to amino acids 374 or 463 (Figure 6A and 7A), possibly deletion of important amino acid sequences located on the N-terminus region could have resulted in improper folding. Generally, balances of different forces are one of the most important factors that contribute to proper folding of proteins and the stabilization of the native protein structure (Fersht 1993; Lazaridis et al. 1995; Pace et al. 1996).

To avoid improper folding issues, another series of NPR1 N-terminus deletion constructs (Figure 8A) were prepared, this time to the full-length NPR1. Subsequently, the amino acid sequence 513 to 535 within the C-terminus region was identified as the potential DNA binding domain that recruits the NPR1 to the -701 fragment of the PR1 promoter in the presence of SA (Figure 8). Our results further support the findings of Wu et al. (2012), who demonstrated that the amino acid sequence 513-593 on the C-terminus is critical for SA and NPR1 interactions, and that the two-cysteine residues 521 and 529 for SA binding are also located within this region (Figure 8A).

Furthermore, we observed that the NPR1 N-terminus deletion constructs (Figure 9A) did not bind to the -870 fragment of the PR1 promoter in the absence of SA (Figure 9B). In addition, another potential DNA binding domain on the N-terminus region of NPR1 (Figure 10 and 11) was identified. Our previous results (see sections 4.2, 4.4 and 4.5), which showed that NPR1 binds to two different sites on the PR1 promoter dependent on the presence or absence of SA, together with the identification of a potential DNA binding domain on the C-terminus amino acid sequence 513 to 535 (Figure 8) that binds to the -701 fragment of the PR1 promoter only in the presence of SA, provided sufficiently convincing evidence that the N-terminus region of NPR1 contains another DNA binding domain for its recruitment to the PR1 promoter in the absence of SA. Through the analysis of a series of NPR1 C-terminus deletion constructs (Figure 10A and 11A) we were able to successfully demonstrate that the amino acids 110 to 190 within the BTB/POZ domain contain a potential DNA binding domain site that recruits NPR1 to the -870 fragment of the PR1 promoter in the absence of SA (Figure 10B and 11B).

These results are also in agreement with the findings obtained by Wu and co-

workers (2012), who have demonstrated that the C-terminus domain of NPR1 is mainly associated with SA-induced interactions, while the N-terminus BTB/POZ domain does not interact with SA with the same affinity as the C-terminus domain in equilibrium dialysis binding assays.

To date, there have been significant insights into the protein-protein interactions that occur in the POZ domain of zinc finger proteins (Aravind and Koonin 1999; Bardwell and Treisman 1994; Boyle et al., 2009; Collins et al., 2001; Li et al., 1999; Rochon et al., 2006). However, considerably less is currently known about the DNA binding properties of the POZ domain (Pessler and Hernandez, 2003). Our obtained results do not come as a surprise; instead, they support other reports on the importance of the POZ domain in promoter recognition and DNA binding activity. For example, Pessler and Hernandez (2003) showed that the FBI-1 POZ domain (Factor that binds to inducer of short transcripts protein 1) specifically binds to inverted sequence repeats downstream of the HIV-1 transcription start site, regulating HIV-1 repression and transcription. Additionally, Katsani et al., (1999) showed that the N-terminus POZ domain of the GAGA protein regulates the formation of oligomers both in vitro and in vivo, and that GAGA oligomers bind DNA with high affinity and specificity. The GAGA transcription factor modulates a variety of stress-induced genes in *Drosophila* (Granok et al., 1995; Wilkins and Lis, 1997).

The last 13 amino acids on the C-terminus of NPR1 (amino acid 580 to 593) were also shown to not participate in the binding of NPR1 to the PR1 promoter in the presence of SA; an investigation into the N-terminus deletion constructs NPR1 463-580 (NPR1 Δ1-463 and Δ580-593) (Figure 7A) and NPR1 1-580 (NPR1 Δ580-593) (Figure 8A) showed that significant reductions in relative *PR1* enrichment values occurred for these constructs after SA addition (Figure 7B and 8B). Amino acid

composition analysis revealed the presence of a stretch of arginine residue (R) in this region. Since arginine could favour the interaction between DNA and protein (Luscombe and Thornton, 2002; Peterson et al., 1996), these results came as a surprise.

Considering these findings, we hypothesized that perhaps amino acids 580 to 593 (NPR1 1-580 (NPR1 Δ580-593)) contain a minimal region involved in intra-molecular interactions with SA that inhibits binding of NPR1 to the PR1 promoter in the presence of SA, which would instead cause a turnover between the domains from (NPR1 513-593 (NPR1 Δ1-513)) to (NPR1 1-190 (NPR1 Δ190-593)), and subsequently, recruitment of NPR1 to the PR1 promoter from SA to non-SA conditions (Figure 18). Consequently, this reaction would switch the NPR1 mode of action on the PR1 promoter from SA to non-SA conditions by folding back (conformational change) and blocking the DNA binding activity of NPR1 513-593 (NPR1 Δ1-513)-701PR1 promoter contacts to facilitate the DNA binding activity of NPR1 (NPR1 Δ190-593)- 870 PR1 promoter contacts (Figure 18). The turnover of amino acids 580 to 593 between NPR1 DNA binding domains from SA to non-SA conditions or vice versa is postulated to happen due to multimerization effects. Mutations in Cys ^{82} and Cys ^{216} resulted in constitutive NPR1 monomerization and *PR1* gene activation (Tada et al., 2008). Cys^{82} is present in the location of identified DNA binding domain of NPR1 (amino acid 110-190) for its recruitment to PR1 promoter -870 fragment in absence of SA. Therefore it can be postulated that in absence of SA, Cys^{82} caused NPR1 oligomerization on PR1 promoter and *PR1* gene repression. This postulation is in agreement with Wu et al., (2012) findings on NPR1 oligomerization. They used *in vivo* 3C experiments and demonstrated that NPR1

oligomer is recruited on PR1 promoter in absence of SA around -734bp to -833bp fragment.

Our results and hypothesis support previous works conducted by Rochon et al. (2006) and Boyle et al. (2009). In their work, the presence of an auto-inhibitory domain on NPR1 was hypothesized to exist, preventing its transactivation function. To support this hypothesis, they showed that NPR1 513-593 (NPR1 Δ1-513) was able to activate transcription in the presence of SA, while the full-length NPR1 protein was not. Additionally, Wu et al. (2012) showed that the binding of SA led to the disruption of the interaction between the BTB/POZ and the C-terminus Δ513 transactivation domain. This resulted in the release of the C-terminal transactivation domain from the auto-inhibition mechanism by the BTB/POZ domain, therefore converting NPR1 into an activated transcription co-activator (see chapter 2, Figure 3).

Similar to our results, intra-molecular interactions have also been associated in the regulation of DNA binding for many transcription factors. A well-explained examples can be found in various works conducted on the Ets-1 (Erythroblastosis Virus E26 Oncogene Homolog 1) protein; for this protein, a loss of connection between the C and N terminal domains has been shown to lead to a conformation in the N-terminal inhibitory region, allowing the C-terminus domain to make stable contacts with DNA (Hagman and Grosschedl, 1992; Jonsen et al., 1996; Lim et al., 1992).

5.7 Importance of Cysteines 521/529 and Cysteine 150 for NPR1 binding to the PR1 promoter in the presence and absence of SA, respectively

EDTA was chosen as a metal chelator agent for experiments seeking to determine whether the binding of NPR1 to the PR1 promoter in the absence of SA is metal-dependent. Our results confirmed the metal-dependency of NPR1 for non SA-induced interactions, as the binding of the NPR1 N-terminal region amino acid 1 to 190 (POZ domain) to the -870 fragment of the PR1 promoter was observed to be completely abolished (Figure 11C). Our findings further support the results obtained by Wu et al. (2012); in their work, the recruitment of SA by NPR1 was shown to be EDTA-sensitive, and its presence as a chelator agent was shown to abolish the binding of SA to NPR1. Additionally, our results are in agreement with the findings of Zschiesche et al. (2015), which demonstrated that the binding of *Arabidopsis* heavy metal-associated isoprenylated plant proteins3 (HIPP3) to the promoter is also metal dependent. They confirmed that the HIPP3 protein is a zinc nuclear binding protein and associated with stress responses. Particularly, regulation of genes involved in the salicylate pathway was affected when the HIPP3 was overexpressed in transgenic *Arabidopsis* plants.

The two cysteines (Cys^{521} and Cys^{529}) on the C-terminus of NPR1 have been shown to be conserved between our identified DNA binding domain for the NPR1 binding to the -701 fragment of the PR1 promoter in the presence of SA and the identified region for SA binding to NPR1 for regulation of *PR1* expression in vivo (Rochon et al., 2006; Wu et al., 2012). In conjunction, these findings point towards Cys^{521} and Cys^{529} as potential metal co-factors for the NPR1 C-terminus region binding to the -701 fragment of the PR1 promoter in the presence of SA (Figure 8).

An analysis of the amino acid composition of the NPR1 N-terminus region, particularly amino acids 110-190, revealed the existence of four cysteine residues (Cys^{150}, Cys^{155}, Cys^{156}, and Cys^{160}) that may act as potential electronegative amino acids for NPR1 binding to the -870 fragment of the PR1 promoter in the absence of SA. The amino acid cysteine (Cys) is involved in many metal protein interactions with DNA (Lando et al., 2000; Toledano and Leonard 1991). As demonstrated by Cao et al. (1997), the Cys^{150} in the N-terminal region is critical for complete NPR1 function and SAR activation. Bearing this in mind, we set out to investigate whether Cys^{150} is also required for metal coordination of the binding of NPR1 to the -870 fragment of the PR1 promoter in the absence of SA. For this determination, Cys^{150}, on the N-terminal region of NPR1, was mutated to Tyr^{150} (Tyrosine 150) by site-directed mutagenesis (Figure 11A). The mutated NPR1 was then submitted to modified ChIP analysis and subsequently, relative *PR1* enrichment analysis. Our results revealed that NPR1-POZ binding to the -870 fragment of PR1 promoter was completely abolished, similar to our obtained results following EDTA treatment (Figure 11D). Hence, we can confirm that in the absence of SA, NPR1 binds to the -870 fragment of the PR1 promoter through metal transition, and that Cys^{150} plays an essential part in the metal-DNA binding of NPR1.

These results also support previously reported findings suggesting the importance of cysteine residues on the NPR1 N-terminus region for SAR-mediated signaling pathways. For example, Mou et al. (2003), through site-directed mutagenesis analysis, demonstrated that the Cys^{82}, Cys^{150}, Cys^{155}, Cys^{160}, and Cys^{216} residues in the N-terminal region are involved with NPR1 oligomer formation and monomer exchange. In the absence of SA, NPR1 is olgomerized via Cys^{82} and Cys^{216} residues and formation of disulfide bonds. Upon SA accumulation and SAR

induction, NPR1 is reduced and returns back to its monomeric form to induce *PR1* expression. Mutation of either Cys^{82} or Cys^{216} residues resulted in constitutive monomerization of NPR1 and *PR1* expression (Mou et al., 2003). In other work, Tada et al. (2008) further established that the S-nitrosylation of NPR1 by S-nitrosoglutathione (GSNO) at Cys^{156} assists in NPR1 oligomerization, which maintains protein homeostasis upon SA induction. Work conducted by Cao and co-workers (1997) revealed that conversion of the cysteine 150 residue to tyrosine 150 in *npr1-2* mutant alleles resulted in *PR1* gene expression reduction, and susceptibility to virulent pathogens, even after SAR induction.

Additionally, through amino acid composition analysis, the DNA binding domain of NPR1 from amino acids 110-190, located in the N-terminal region, was revealed to contain negatively charged acidic residues (E), polar uncharged (C, Q, S), non-polar aliphatic (G, L), and positively charged basic residues (H). These findings are in agreement with work conducted by Gupta and co-workers (2012), in which rice resistance protein Pi45, containing two zinc finger domains, was revealed to also hold a similar amino acid sequence composition within its domains. These charged amino acids (E and H) have been shown to contribute to electrostatic interactions that provide essential force for the stability of protein-DNA interactions (Kumar et al., 2008). As such, it can be assumed that the identified negatively-charged acidic residues participate to stabilize the interaction between NPR1 and the -870 fragment of the PR1 promoter that occurs in the absence of SA.

Our results represent a novel advancement on plant SAR-mediated disease resistance by confirming the direct binding of NPR1 to the PR1 promoter. Figure 18 presents a model that summarizes our results on the regulation of the *PR1* gene.

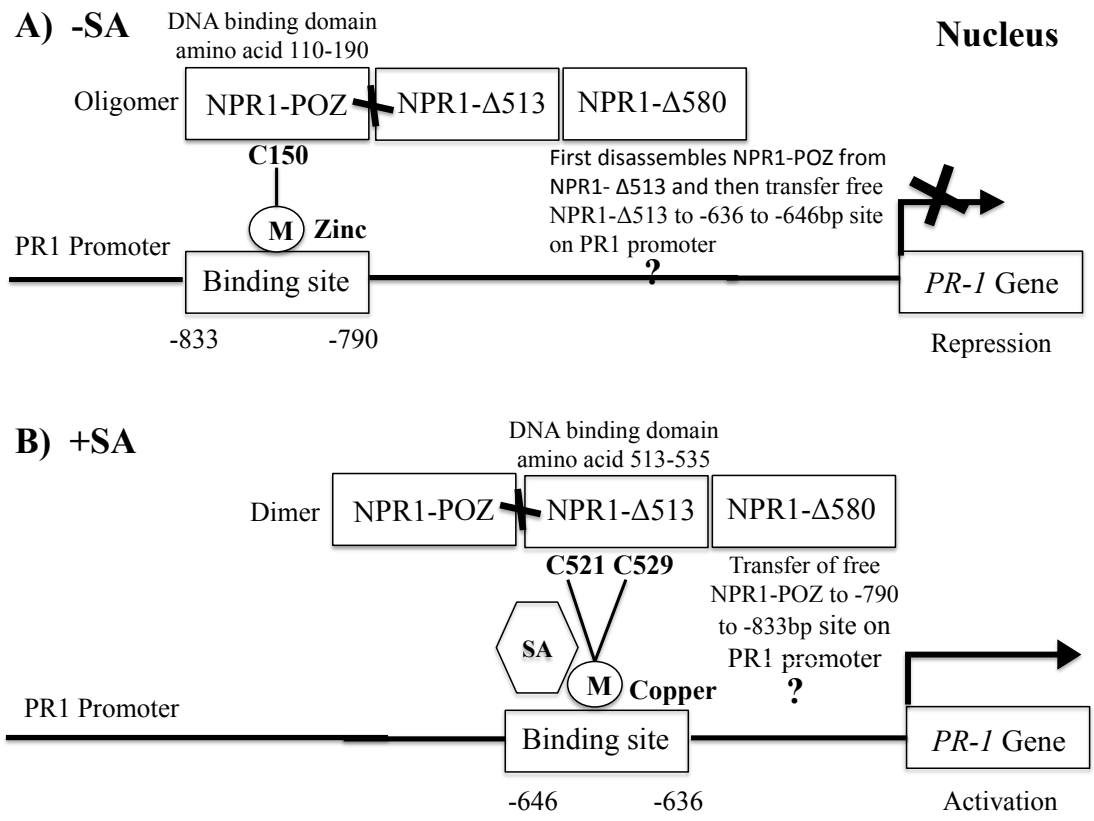


Figure 18. The model for regulation of *PR1* by direct binding of NPR1 to the PR1 promoter

A. Scheme showing that in the absence of SA and under resting conditions, NPR1, via its DNA binding domain located in the N-terminus region, POZ domain (amino acid 110-190) and with the assistance of the transition metal zinc (data not shown), is directly recruited to the PR1 promoter, specifically -790 to -815 base pair fragment. Cysteine 150 residue is critical for coordination with zinc metal (data not shown). The *PR1* gene is repressed at this state. NPR1-Δ580 is postulated to be involved in a switch mechanism from -SA to +SA condition and vice versa. In –SA condition NPR1-Δ580 first disassembles the NPR1-POZ domain from NPR1- Δ513 domain and then transfer free NPR1- Δ513 for interaction with PR1 promoter (-636bp to -646bp). However, the exact nature of this mechanism remains undetermined. The question mark illustrates this fact. The turnover of NPR1-Δ580 between NPR1 DNA binding domains from SA to non-SA conditions or vice versa is postulated to happen due to multimerization effects. Mutations in Cys ^{82} and Cys ^{216} resulted in constitutive NPR1 monomerization and *PR1* gene activation (Mou et al., 2003). Cys^{82} is present in the location of identified DNA binding domain of NPR1 (amino acid 110-190) for its recruitment to PR1 promoter -870 fragment in absence of SA. Therefore in absence of SA, Cys^{82} could cause NPR1 oligomerization on PR1 promoter and *PR1* gene repression.
B. Scheme demonstrating the mechanism that results in activation of the *PR1* gene in the presence of SA. NPR1 is recruited to the PR1 promoter -636 to -646 base pair fragment through its DNA binding domain located in the C-terminus region NPR1- Δ513 domain (amino acids 513-535). Cysteine 521 and 529 residues participate in NPR1 metal (copper) binding and direct SA perception. As demonstrated by Wu et al. (2012), the binding of SA leads to the disruption of the interaction between the BTB/POZ and the C-terminus

(NPR1- Δ513), thus releasing the C-terminal transactivation domain from auto-inhibition by the BTB/POZ domain. Therefore, the C-terminus (NPR1- Δ513) is free to directly bind to the PR1 promoter and *PR1* gene activation. NPR1-Δ580 is postulated to be involved in a switch mechanism from -SA to +SA condition and vice versa. In +SA condition, the NPR1-POZ domain is released from NPR1- Δ513 domain due to conformational changes caused after SA perception by NPR1- Δ513 domain. So it is postulated that NPR1-Δ580 transfer the free NPR1-POZ to other site of PR1 promoter (-790 to -833bp) for interaction with PR1 promoter. However, the exact nature of this mechanism remains undetermined. The question mark illustrates this fact.

5.8 Competition of 4-OH-BA with SA and inhibition of NPR1 binding to the -701 fragment of the PR1 promoter

In a previous section (see section 5.4), we made note that relative *PR1* enrichment was observed to be abolished for the -701 and -870 fragments of the PR1 promoter in the presence of 4-OH-BA (Figure 3D). These results are in agreement with previously published reports (Wu et al., 2012) that maintain that 4-OH-BA competes with SA for NPR1 binding, which prompted our hypothesis that 4-OH-BA might act as a DNA binding inhibitor by competing with SA for NPR1 binding to the PR1 promoter. To test this hypothesis, 4-OH-BA was incubated from low to high concentrations (1μM, 10μM, and 80μM) with two SA concentrations (1μM and 10μM), which were kept constant. The results from this set of experiments revealed that 4-OH-BA competes with SA for NPR1 binding, inhibiting NPR1 recruitment to the -701 fragment of the PR1 promoter (Figure 13C and D). However, no statistically significant differences were observed between the 1, 10, and 80μM concentrations of 4-OH-BA, suggesting that the binding sites on NPR1 were already saturated at lower concentrations of 4-OH-BA; as such, increases in concentration did not contribute to further drastic changes in NPR1 binding.

To further explore this line of inquiry, NPR1 513-593 (NPR1 Δ1-513) and NPR1 533-593 (NPR1 Δ1-533) were incubated with a series of high to low concentrations of 4-OH-BA (1, 0.1, 0.00001, 0.000001, 0.0000001μM). As 4-OH-BA concentrations decreased, the amounts of relative *PR1* enrichment obtained for NPR1 513-593 (NPR1 Δ1-513) and NPR1 533-593 (NPR1 Δ1-533) were observed to correspondingly increase. Remarkably, 4-OH-BA was demonstrated to have a significant effect on the binding of NPR1 even at a very low concentration (0.0000001μM) (Figure 14B and C), suggesting that even small quantities of 4-OH-

BA are enough to produce an inhibitory reaction. These results clearly demonstrated that 4-OH-BA acted antagonistic of SA and inhibited the binding of NPR1 to the -701 fragment of the PR1 promoter.

Notably, these results further support both our previous findings as well as the reported findings of Wu et al. (2012) regarding the relationship between 4-OH-BA and SA. Through our findings (see sections 4.7 and 5.6), NPR1 513-593 (NPR1 Δ1-513) and NPR1 533-593 (NPR1 Δ1-533) are further confirmed as the specific sites on the NPR1 C-terminus region for SA perception (Wu et al., 2012). Further, our findings establish that the DNA binding domain that facilitates NPR1 recruitment to the -701 fragment of the PR1 promoter in the presence of SA is localized between amino acids 513 to 535 on the C-terminus region of NPR1 (Figure 8).

The identification of 4-OH-BA as a potent DNA inhibitor does not come as a surprise; several reports have been published on the use of plant extracts as potent DNA-binding inhibitors. For example, Diterpenoids, isolated from the Chinese herb *I. rubescens*, have been shown to inhibit the DNA-binding activity of the nuclear factor-κB (NF- κB) transcription factor. The transcription induced by nuclear factor-κB (NF-κB) in mammals is highly similar to the one induced by NPR1 in plants (Leung et al., 2005).

Through in vivo experiments, we were further able to demonstrate that in natural conditions (i.e., living plants), 4-OH-BA inhibits NPR1 binding to the PR1 promoter, and ultimately, *PR1* gene activation (Figure 15). In these experiments, upon treatment of plants with both SA and 4-OH-BA, 4-OH-BA was demonstrated to regulate *PR1* transcription antagonistic of SA by competing with SA. Additionally, 4-OH-BA was shown to negatively regulate *PR1* transcripts, opposed to SA, at higher concentrations (3 or 30μM), while at lower concentration (0.3μM) the *PR1* transcripts

were observed to return back to the SA-induced levels (Figure 15B). Although other inactive SA analogues have similar structures to 4-OH-BA, plants sprayed with combinations of SA and either 3-OH-BA and BA, using different concentrations (30, 3, and 0.3μM), were observed to yield less *PR1* transcripts at all concentrations (Figure S4 B and C).

Among all analogues of 4-OH-BA at higher concentrations (30μM), only catechol (Figure 15A) showed a higher abundance of *PR1* transcripts similar to SA-induced leaves. Conversely, at lower concentrations (0.3μM), lower *PR1* transcripts were obtained (Figure 15C), suggesting that catechol competed synergistically with SA in a manner dissimilar to 4-OH-BA for NPR1 binding to the PR1 promoter, and ultimately, *PR1* expression. Our obtained results, which have shown that in comparison to 4-OH-BA, catechol induces an opposite response on *PR1* expression, also confirmed that the antagonistic effect of 4-OH-BA, as well as its competition with SA for NPR1 binding to the PR1 promoter, are not generic effects observed among all benzoic acid family members; rather, these are specific features of 4-OH-BA. These results are also in agreement with the findings obtained by Friedrich and co-workers (1995) who have observed that upon catechol infiltration in tobacco plants, lesions formed by tobacco mosaic virus infection remained unchanged. However, when catechol was co-infiltrated with different SA concentrations, *PR1* expression was not reduced or changed.

Once the role of 4-OH-BA was confirmed in vitro and in vivo as a potent DNA inhibitor of the NPR1-PR1 promoter binding, an experimental procedure was set up for further in vivo confirmation of the observed in vitro findings. For this purpose, the abundance of *PR1* transcripts in wild-type and *PAL* -mutant *Arabidopsis* plants were observed prior to and after treatment with SA. PAL –mutant *Arabidopsis*

plants were chosen as samples due to the unique nature of their mutation; as shown in the literature, the 4-OH-BA is formed through the phenylpropanoid pathway (Löscher and Heide, 1994) and also directly from chorismate by chorismate pyruvate-lyase (CPL) (Kohle et al., 2002; Meganathan, 2001). In the phenylpropanoid pathway, phenylalanine ammonia-lyase (PAL) is activated upon pathogen challenge, converting phenylalanine (Phe) to cinnamic acid (Bagal et al., 2012). The cinnamic acid is then converted to 4-OH-BA (Kohle et al., 2002; Smith-Becker et al., 1998). Therefore, in *PAL* -mutant *Arabidopsis* plants, 4-OH-BA synthesis is altered but not completely abolished since the CPL enzyme is still present for 4-OH-BA production.

Our results revealed that a higher degree of *PR1* transcripts occurred for *PAL* mutant plants compared to SA-induced wild-type plants (Figure S4 E); however, differing levels of *PR1* transcripts were obtained for *PAL1* and *PAL2* (two different alleles of *PAL* mutant). As such, we were unable to conclude that the inhibitory function of 4-OH-BA in NPR1 DNA-binding was completely abolished in *PAL* mutant plants due to loss of 4-OH-BA production. We expected to see a similar induction of *PR1* transcripts in SA-treated PAL mutants as wild-type plants. If 4-OH-BA synthesis had been observed to be completely abolished in PAL mutant plants, we would have been able to conclude that no competition occurred between SA and 4-OH-BA, and that therefore, SA binding to NPR1 and subsequent binding to PR1 promoter could result in *PR1* expression in PAL mutant plants, similar to wild-type plants.

An explanation for the inconsistencies observed between levels of *PR1* transcripts for the two alleles of *PAL* mutant (*PAL1* and *PAL2*) can perhaps be found in the role played by chorismate pyruvate-lyase (CPL), an enzyme involved in 4-OH-BA synthesis (Kohle et al., 2002). It is possible that CPL might have contributed in 4-

OH-BA synthesis in *PAL* mutant plants, which might have affected SA accumulation and subsequent *PR1* transcripts.

It is also important to note that mixed reports have been published on the importance of PAL activity, 4-OH-BA synthesis, and SA accumulation. For example, *PAL* activity has been shown to be essential for SAR induction and SA accumulation both in tobacco and *Arabidopsis* plants (Mauch-Mani and Slusarenko 1996; Pallas et al., 1996). Transient increase in PAL activity was detected in *Arabidopsis* upon elicitor treatment followed by an increase in 4-OH-BA (Schnitzler and Seitz, 1989; Bach et al., 1993). Contrastingly, Smith-Becker et al., (1998) showed that after pathogen inoculation in cucumber plants and increase in PAL activity, 4-OH-BA accumulated in phloem fluids with the same kinetic as SA, and in contrast to SA, 4-OH-BA infiltration did not result in pathogen resistance, nor in the induction of SA accumulation in phloem exudates.

In addition, the hydroxyl (OH) group is a common factor between 4-OH-BA and 3-OH-BA. Indeed, *PR1* transcripts levels were observed to be similar for plants sprayed with either compound. However, in the case of catechol, *PR1* transcripts results were observed to be different to the results of 4-OH-BA, confirming that catechol behaved completely opposite of 4-OH-BA. Within its molecular structure, catechol hydroxyl (OH) group is in adjacent position but in BA, 3-OH-BA, and 4-OH-BA the hydroxyl groups are not in adjacent positions. Together with our findings, the molecular structure of catechol serves to further suggest that the hydroxyl group adjacent position is associated with the function of 4-OH-BA as an inhibitor of NPR1 binding to the PR1 promoter.

5.9 Binding of additional members of the NPR family, NPR3, NPR2, and NPR4, to the PR1 promoter

Five other *NPR1-like* genes have been previously identified in the *Arabidopsis* genome, including *NPR2*, *NPR3, NPR4, NPR5* (*BLADE-ON-PETIOLE2*), and *NPR6* (*BLADE-ON-PETIOLE1*) (Figure 5SA) (Hepworth et al., 2005; Liu et al., 2005).

In the currently presented work, we were able to demonstrate that NPR2, NPR3, and NPR4 are also recruited to the PR1 promoter. Similar to NPR1, NPR4 was also recruited to the -701 fragment only in the presence of SA, and to the -870 fragment only in the absence of SA (Figure 16B and C). On the other hand, NPR3 was not recruited to the -701 fragment, neither in the presence nor in the absence of SA; however, it was shown to be recruited to the -870 fragment in the absence of SA (Figure 16B and C).

Although NPR3 and NPR4 share 34.5% and 36% sequence similarity with NPR1, respectively (Table S1), and despite the fact that all three have been shown to contain four conserved cysteine residues in their BTB/POZ domain in addition to a stretch of 5 different basic amino acids in their C-terminus domain (Figure 16A) (Liu et al., 2005; Shi et al., 2013), our results clearly demonstrate that they each have different binding profiles to the PR1 promoter, both in the presence and absence of SA. Consequently, a high amino acid sequence similarity between proteins may not always be indicative that they share similar functions (Pearson, 2013).

Our findings revealed that NPR1 and NPR4 have similar binding profiles to the PR1 promoter in the presence and in the absence of SA. This prompted the hypothesis that similar to NPR1, NPR4 could also act as a positive regulator of SAR. However, our results do not support the model proposed by Fu and co-workers (2012). In this model, both NPR3 and NPR4 are hypothesized to be possible SA

receptors, similar to NPR1; NPR4 acts as a CU13 E3 ligase adopter in the absence of pathogen challenge, targets NPR1 for degradation, and continues turnover by proteasome. However, after SA induction, SA binds NPR4 and abolishes the interaction between NPR4 and NPR1. When SA is accumulated and plants are challenged with pathogen, NPR3, instead of NPR4, acts as a CU13 E3 ligase adopter. Based on this model, both NPR3 and NPR4 play roles as substrates and SA receptors, which allows for the degradation of NPR1 in SAR-induced and naïve cell states, respectively. However, our results have shown that both NPR1 and NPR4 bind to the -701 fragment of the PR1 promoter in the presence of SA, while in the absence of SA, they bind to the -870 fragment. Consequently, NPR4 cannot act as a CU13 E3 ligase adopter and degrade NPR1. Instead, our results suggest that NPR4, similar to NPR1, might be involved in the early stages (12hrs post SA induction) of SAR and *PR1* expression, when SA is present. Similar to NPR4, our results also suggest that NPR3 cannot act as a CU13 E3 ligase adopter, targeting NPR1 for degradation when SA is accumulated. As our results revealed, in the presence of SA, the binding of NPR3 to the -701 fragment was completely abolished. As a result, it can be assumed that NPR3 is only involved in *PR1* enrichment and SAR induction in the absence of SA. In theory, NPR3 could also play a role as a positive regulator of SAR in later stages (24 hrs post SA induction), when SA induction is not needed. These results are also in agreement with the finding obtained by Fu and co-workers (2012) because they also demonstrated that NPR3 showed lower affinity for SA binding compared with NPR4.

Chiefly, our results are in agreement with the reported findings of Kuai et al. (2015) and Zhang et al. (2006). Both groups have suggested that NPR3 and NPR4 play a redundant function in plant immune regulation, and that NPR3 and NPR4 are

both SA-dependent receptors at high (early *PR1* activation) and low (late *PR1* activation) SA concentrations.

Contrary to some of the relevant literature, our results have demonstrated that similar to NPR1, both NPR3 and NPR4 positively regulate SAR. The source of some of the discrepancies reported between Liu et al. (2005), Zhang et al. (2006), and Fu et al. (2012) most likely arise from the use of different *npr1/npr3/npr4* mutant alleles in experiments. As evidenced in their research, the *npr1/npr3/npr4* mutants presented different phenotypes across different reports. For example, Fu et al., (2012) showed that single *npr3* or *npr4* mutant plants were similar to wild-type plants in disease resistance response. However, both Zhange et al., (2006) and Liu et al., (2006) demonstrated that *npr4-3*, *npr4-2*, *npr4-1* mutant plants were not similarly resistant to bacterial pathogens as wild-type plants.

Particularly, a number of concerns have been raised in regards to the model proposed by Fu et al. (2012). For example, the interaction between NPR3 and NPR4 with SA has fallen under scrutiny, as the authors failed to demonstrate the proposed binding using equilibrium associated binding assay; instead, the yeast two-hybrid method was employed to characterize this binding. As well, as typical characteristics of any hormonal receptor interaction, the conformational changes for NPR3 and NPR4 in the presence of SA remained to be demonstrated. They only showed that NPR4 was presented in tetramer form both before and after SA and addition of SA did not change the elution profile of NPR4. In addition, NPR1 has since been confirmed to bind to SA (Manohar et al., 2015; Wu et al., (2012); in the proposed model, the interaction between SA and NPR1 was not demonstrated either. Lastly, as demonstrated by Wu et al. (2012), NPR1 has been shown to undergo conformational change after SA binding; it is likely, accordingly, that NPR1 is the decisive factor in

whether any interactions occur between NPR3 and NPR4, regardless of the presence or absence of SA (Kuai et al., 2015).

Although NPR2 shows the highest similarity (61.89%) with NPR1 (Table S1), little is known about its function. In the current work, NPR2 has been demonstrated to bind to the -870 fragment of the PR1 promoter in the absence of SA, although it does not undergo binding to the -701 fragment in the presence of SA (Figure 16B and C). Our results support the findings reported by Canet et al. (2010), which show that NPR2 is involved in the SA signal transduction pathway for plant defense regulation. Our results further confirmed that NPR2 is associated with defense responses, but only in the absence of SA. In the presence of SA, binding of NPR2 to the PR1 promoter was observed to be completely abolished. Despite the high degree of similarity between NPR2 and NPR1, our results indicate that NPR2 may play a different role as NPR1 in defense signaling pathway.

Our results further indicated that neither NPR5 nor NPR6 recruit to the PR1 promoter, in the presence or absence of SA (Figure 16B and C). These results further validate other reported findings on NPR5 and NPR6 functions, which propose that both NPR5 and NPR6 are not associated with plant defense responses; instead, they are likely involved with developmental processes within plants (Ha et al., 2004; Hepworth et al., 2005; Norberg et al., 2005; McKim et al., 2008; Jun et al., 2010; Saleh et al., 2011).

Amino acid composition analysis (Figure S6) showed that Cys^{521} and Cys^{529}, the two cysteine residues critical for SA perception and NPR1 binding to the -701 fragment of the PR1 promoter, are not conserved in NPR2, NPR3, and NPR4. In NPR2 and NPR3, the amino acids aspartic acid (D) and glutamic acid (E) are positioned in place of Cys^{521} and Cys^{529}, respectively. Considering that both NPR2

and NPR3 only bound to the -870 fragment of the PR1 promoter in the absence of SA, this would suggest that these amino acids are not critical for SA perception by the NPR protein family. In NPR4, the amino acids proline (P) and arginine (R) are positioned in place of Cys^{521} and Cys^{529}, respectively. As mentioned earlier, NPR4 showed similar binding profiles as NPR1, being recruited to the -701 and -870 fragments of PR1 promoter in the presence and absence of SA, respectively. Accordingly, in addition to cysteine residues, it is plausible that proline and arginine residues are also important factors in SA perception by NPR proteins. However, these hypotheses still remain to be tested.

5.10 Binding of maize and rice NPR1 proteins to the PR1 promoter via SA and metal co-factors

In the current work, we tested the recruitment of corn and rice NPR1 proteins to both the -701 and -870 fragments of the PR1 promoter, in the absence and presence of SA. The obtained results indicated that both corn and rice NPR1 proteins are recruited to the -701 fragment of PR1 promoter in the presence of SA, while neither bind to the -870 fragment of the PR1 promoter. Additionally, both rice and corn NPR1 were observed to be sensitive to EDTA treatment, suggesting that similar to *Arabidopsis* NPR1*,* their bindings to the PR1 promoter may occur through a metal co-factor (Figure 17).

Notably, our findings add to the evidence that SA-mediated defense responses through NPR1 are conserved among plant species. Furthermore, our results support other published reports on SAR-mediated defense pathways of rice and corn. For example, Rostas and Turlings, (2008) observed that in maize (corn) leaves, fungal infection resulted in SA accumulation as well as exogenous application of SA-suppressed disease symptoms. In other studies, overexpression of *Arabidopsis* NPR1 in transgenic rice plants was shown to enhance disease resistance against fungal pathogens (Fitzgerald et al. 2004; Quilis et al. 2008). Overexpression of rice NPR1 also enhanced disease symptoms, while its knockdown resulted in increased susceptibility to infection (Chern et al. 2005; Yuan et al. 2007).

Wu et al. (2012) reported that the Cys^{521} and Cys^{529} residues responsible for SA binding to *Arabidopsis* NPR1 are not universally conserved in other NPR1 from different plant species, such as rice and corn. As well, Silverman et al. (1995) demonstrated that rice has a higher basal SA level (unchallenged condition) as compared to tobacco and *Arabidopsis*, leading the authors to hypothesize that rice

NPR1 might have different interaction mechanisms with SA. However, as evidenced by the current work, rice NPR1 has been shown to interact with SA in a manner similar to *Arabidopsis* NPR1, through a metal co-factor. Instead, the obtained results support the findings of Xiao and Wedd (2010); in their work, the authors demonstrated that metal-protein interactions are not restricted to the cysteine residues. Further, the authors maintained that any amino acid that holds electronegative elements in its side chain has the potential to interact with metals. An analysis of the amino acid compositions (Figure S7) and phylogenetic relationship (Table S2 and Figure S5B) of rice and corn NPR1s revealed that a low percentage of sequence identity exists among the two plants. Additionally, both rice and corn NPR1s were shown to exclude a C529 equivalent; however, an aspartic acid amino acid carrying a carboxylate in its side chain (electronegative atom) presented around the position where C521 would be expected to be found. As such, it can be hypothesized that rice and corn are able to interact with metals and consequently, an assumption could be made that in different crops, SA-NPR1 interactions occurring via metal co-factor would share a degree of similarity from species to species.

Moreover, our results on the binding of maize NPR1 to the PR1 promoter are in agreement with work conducted by Chen et al. (2012). In our research, *Zm*NPR1 was shown to be recruited to the PR1 promoter only in the presence of SA and through a metal co-factor (Figure 17). In their work, Chen and co-workers (2012) demonstrated that expression levels of NPR1 from sugarcane (*Sc*NPR1) (*Saccharum spp*.) were induced after SA treatment and pathogen inoculation. In addition, the SA-mediated defense signaling pathways were shown to also be conserved in sugarcane plants; phylogenetic analysis results showed that *Sc*NPR1 was significantly similar to NPR1 from maize (83% sequence similarity), and that both share common

characteristics with other NPR1 proteins from different plant species. For example, the authors reported that *At*NPR1 and *Zm*NPR1 share a common BTB zinc finger domain and an ankyrin repeat. Additionally, our results support the findings of Balmer et al. (2013) that showed SAR in maize plants is associated with accumulation of SA both at local and distal tissues. These findings, together with our own, suggest that *Zm*NPR1 is also involved in SA-mediated signaling responses associated with increases in resistance levels against pathogen infection.

Isotianil is an SA analog previously shown to enhance disease responses against fungal pathogens and induce the same components of the SAR pathway as SA in rice plants (Brozek et al., 2012; Ogava et al., 2011; Yoshida et al., 2013). Accordingly, in the current work, Isotianil was tested on corn, rice, and *Arabidopsis* NPR1 proteins for binding to the PR1 promoter. However, at this time, a determination could not be made as to whether Isotianil acts as SA and induces SAR in *Arabidopsis*, rice, and corn (Figure S8). As a possible explanation, the low solubility of Isotianil in the tested solvents could have played a factor in obtaining accurate results, as the Isotianil chemical is manufactured in a way that they stay intact, particularly when sprayed outside in a field environment.

CHAPTER 6- CONCLUSION AND FUTURE PERSPECTIVES

6.1 Overall conclusions of the study

In the herein presented research, a modified in vitro ChIP method was successfully validated and applied towards the determination of *Arabidopsis thaliana* NPR1 as a novel DNA binding protein that is recruited to the PR1 promoter, similarly to the TGA2 transcription factor.

NPR1 was shown to bind to the PR1 promoter specifically to the -636 to -646 base pair sequence (LS7 site) in the presence of SA and to the -790 to -833 base pair fragment in the absence of SA. Moreover, two distinct DNA binding domains were identified within both the C and N-terminus regions of NPR1. Particularly, our results showed that the DNA binding domain that binds to the -701 fragment of the PR1 promoter in the presence of SA was localized between amino acids 513 to 535, and that the Cysteine 521 and 529 residues, as well as the transition metal copper, play a critical role in the binding of SA to NPR1, and the subsequent recruitment of NPR1 to the PR1 promoter. Similarly, the DNA binding domain of NPR1 for binding to the -870 fragment of the PR1 promoter in the absence of SA was revealed to be located between amino acids 110 to 190, with the transition metal and the Cysteine 150 residue shown to be required for this interaction to occur. Similar to any protein DNA interaction recognition mechanism, our results showed that NPR1 only binds to the PR1 promoter when plasmid DNA is present in circular, double-stranded format.

The use of Mfold DNA secondary structure prediction software allowed for the postulation that the formation of a Stem-loop (hairpin) and multi-loops around the

binding sites of NPR1 on the PR1 promoter contribute to the stabilization of the interaction complex.

Other inactive analogues of SA, including INA and 4-OH-BA, were revealed to negatively affect the binding of NPR1 to the PR1 promoter. Further in vitro analysis conducted on the inhibitory effect of 4-OH-BA in the binding of NPR1 to the PR1 promoter allowed for the corroboration that 4-OH-BA acts as an antagonist of SA and inhibits the binding of NPR1 to the PR1 promoter via competition with SA for NPR1 binding sites. Further experimentation in vivo demonstrated that this process occurs even at low concentrations. Conversely, at higher concentrations, 4-OH-BA was demonstrated to negatively regulate *PR1* transcription in *Arabidopsis* leaves, as opposed to SA. Among all analogues of 4-OH-BA, at higher concentrations, catechol was found to synergistically compete with SA opposed to 4-OH-BA for NPR1 binding to the PR1 promoter, and ultimately, *PR1* expression.

NPR2, NPR3, and NPR4 were also shown to bind to the PR1 promoter. Our results demonstrated that similar to NPR1, both NPR3 and NPR4 act as positive regulators at different stages of SAR establishment. Conversely, NPR5 and NPR6 were shown to abstain from binding to the PR1 promoter, further confirming they are not associated with plant defense responses. The use of NPR5 and NPR6 proteins as negative controls further confirmed the validity of our modified in vitro ChIP method.

In addition to *Arabidopsis* NPR1, NPR1 proteins from economically relevant crops, such as maize and rice, were shown to also be recruited to the PR1 promoter in the presence of SA and a metal co-factor. These results further support the concept that SA-mediated defense responses through NPR1 are conserved among plant species.

In conjunction, the presented findings expand our understanding of how the

NPR1 protein behaves with the PR1 promoter to regulate *PR1* gene expression during SAR. In addition, this study also revealed the suitability of the modified in vitro ChIP method as a primary screening technique for discovery of novel plant immune-priming compounds (chemical activators) capable of replacing hazardous pesticides, and similar to SA, aid in the protection of economically important crops through NPR1-mediated SAR defense responses.

6.2 Future prospects

Through application of the in vivo ChIP method, Rochon and co-workers (2006) were able to demonstrate that NPR1 was recruited to the PR1 promoter in both resting and SA-treated cells on *Arabidopsis* wild type plants. However, they failed to make final conclusions due to insufficient experimental evidence. Contrastingly, the currently presented work was able to confirm in vitro recruitment of NPR1 to the PR1 promoter. As a future direction, it would be notable to conduct in vivo ChIP experiments on wild-type (positive control) and npr1-mutant (negative control) plants before and after treatment with SA, followed by massive parallel DNA sequencing (ChIP-seq) (see section 2.2.4.1) (Najafabadi et al., 2015). The pursuit of in vivo confirmation of our results in a genome, and within the native chromatin context, would serve to further enrich our understanding of the NPR1-PR1 promoter process, as well as add credence to the currently presented findings.

Through the application of the validated in vitro ChIP method, we were able to show that the DNA binding domains located in the C and N-terminal regions of NPR1 were critical for binding to the PR1 promoter, in the presence and absence of SA, respectively. Future applications of isothermal titration calorimetry (ITC) would serve to further confirm the observed binding activity (see section 2.2.4.3) (Jin et al., 2015; Liu et al., 2015; Ross et al., 2015). Such experiments would provide a full thermodynamic description of the binding of NPR1 to the PR1 promoter, further enriching our understanding of this process. Particularly, the application of ITC would allow for the attainment of quantitative information on the stoichiometries, binding affinity, enthalpy, and entropy of this interaction. Considering that the critical regions for interaction between the NPR1 protein and the PR1 promoter have been successfully localized in the currently presented research, successful application of

ITC analysis would be easily achievable; with our current knowledge, the amino acid and base pair sequences of the regions could be chemically synthesized and used in ITC experiments.

In the absence of TGA2, NPR1 has been shown to directly bind to the PR1 promoter. In vivo results by Rochon et al. (2006) initially demonstrated that TGA2 is recruited to the PR1 promoter in the absence of both SA and NPR1. Further research by Boyle and co-workers (2009) revealed that the NPR1-TGA2 enhancesome on the PR1 promoter plays a critical role in de-repressor activity of TGA2. While investigations into the behavior of TGA2 fell outside of the scope of the current work, it would be of considerable scientific interest to apply a similar modified in vitro ChIP approach to test the binding of NPR1 to the PR1 promoter in the presence of TGA2. This would allow us to gain further insight into the role of TGA2 in the recruitment of NPR1 to the PR1 promoter, and consequently, into the roles of TGA2 and NPR1 in plant resistance responses.

As previously discussed in sections (2.1.5.4.4.3) and (5.9), Fu et al. (2012) proposed a model describing how both NPR3 and NPR4 degrade the NPR1 protein and negatively regulate SAR. Our findings, however, have been shown to contradict this model; in the currently presented work, NPR1 and NPR4 were revealed to display similar binding profiles to the PR1 promoter in the presence and absence of SA, while NPR3 was demonstrated to only be recruited to the PR1 promoter in the absence of SA. Consequently, we hypothesized that both NPR3 and NPR4 could function as positive regulators of SAR, although at different stages of establishment. Contrastingly, the model proposed by Fu et al. (2012) expounds that both NPR3 and NPR4 play roles as substrates and SA receptors, allowing for the degradation of NPR1 in resting and SAR-induced states, respectively. Considering the discrepancies

associated with the proposed model in regards to our results, further experimental data could help clarify the roles of the NPR-family proteins, while helping clear up other concerns that been raised by the scientific community. Fu et al. (2012) were only able to demonstrate the binding of SA to NPR3 and NPR4 using the yeast two-hybrid method, which has been shown in the past to carry some measure of unreliability (Stellberger et al., 2010). Thus, the application of equilibrium dialysis binding assays similar to the approach taken for NPR1 (Wu et al., 2012) would assist in clarifying the results obtained by Fu et al. (2012) in regards to the interaction of SA with NPR3 and NPR4. As well, it has been noted that in the research reported by Fu et al. (2012), the authors failed to include evidence of conformational changes in NPR3 and NPR4 after interaction with SA. As typical characteristics of any hormonal receptor interaction, the effect of SA on conformations of NPR3 and NPR4 could potentially be tested through the use of the gel filtration approach. Additionally, 3C *in vivo* experiments should be performed to find out whether NPR3 or NPR4 binds SA and their stoichiometry changes on PR1 promoter or not. Major discrepancies have also been reported between the works of Liu et al. (2005), Zhang et al. (2006) and Fu et al. (2012) in regards to the phenotypes of the *npr3* and *npr4* mutant plants in response to pathogen challenge. As such, it would be advantageous to monitor *PR1* expression in mutant and wild-type mature *Arabidopsis* plants after pathogen challenge and SA addition with the use of the RT-PCR approach and a reliable reference gene. This would provide information on the roles of NPR3 and NPR4 in SAR-mediated disease responses while helping clear up major discrepancies reported between groups.

As discussed in section (5.8), inconsistent results were obtained for *PR1* transcripts in two different alleles of *PAL*-mutant (*PAL1* and *PAL2*) *Arabidopsis* plants. We postulated that perhaps the chorismate pyruvate-lyase (CPL) enzyme

(Kohle et al., 2002) contributed to 4-OH-BA synthesis in *PAL* mutant plants, which in turn might have affected SA accumulation and subsequent *PR1* transcription. Although the CPL enzyme has not yet been characterized in plants, upon its characterization, it would be imperative to test *PR1* transcripts in CPL enzyme mutants and wild-type plants before and after SA addition. Such experiments would further elucidate the role of 4-OH-BA as a potent inhibitor of NPR1 binding to the PR1 promoter and provide in vivo information on the antagonistic relationship between SA and 4-OH-BA.

Furthermore, we could not confirm the formation of stem and multi-loop secondary structures using classical *S1* endonuclease enzyme mapping assay. In the future, a different approach, such as atomic force microscopy (AFM) (Lavalley et al., 2007), could aid in the confirmation of the predicted structure formations obtained with the use of Mfold software that occur on the PR1 promoter around NPR1 binding sites. AFM experiments would allow for the attainment of topographic images of secondary structures, providing evidence that confirms the contribution of secondary structures to the stabilization of NPR1-PR1 promoter interactions.

NPR1 proteins from rice and maize were shown to also bind to the PR1 promoter in the presence of SA and a metal co-factor. For future research, equilibrium dialysis binding analysis, similar to the approach taken for *Arabidopsis* NPR1 (Wu et al., 2012), could be applied towards these crops, further confirming the direct interaction between SA and NPR1 proteins from maize and rice. In addition, NPR1 proteins from other economically important crops, such as soybean, could be similarly investigated for PR1 promoter interactions in the presence of SA.

Lastly, the modified in *vitro* ChIP method can be used as a primary screening technique in the identification of the DNA binding domains and binding sites of rice

and soybean NPR1 proteins on the PR1 promoter, in addition to aiding in the localization of regions involved with SA interaction. This would allow for chemical synthesis of the localized regions and subsequent testing of their interactions with potential chemical compounds (Noutoshi et al., 2012), which could be accomplished with the use of thermal shift assays as a secondary screening method. As rice and soybean are economically important crops worldwide, the knowledge gained through these experiments could aid in investigations of plant immune-priming compounds (chemical activators). These compounds, in turn, could be used in place of hazardous pesticides in the protection of crops, through the use of NPR1-mediated SAR defense responses.

REFERENCES

Adie, B.A.T., Perez-Perez, J., Perez-Perez, M., M., Godoy M., Sanchez-Serrano J., J., Schmelz, E., A., Solano, R. (2007). ABA is an essential signal for plant resistance to pathogens affecting JA biosynthesis and the activation of defenses in Arabidopsis. *Plant Cell* **19**:1665–1681.

Aich, P., Labiuk, S., L., Tari, L., W., Delbaere, L., J., T., Roesler, W., J., Falk, K., J., Steer, R., P., and Lee, J., S. (1999). M-DNA: A complex between divalent metal ions and DNA, which behaves as a molecular wire. *J. Mol. Biol.* **294**: 477-485.

Alber, T. (1992). Structure of the leucine zipper. *Curr Opin Genet Dev* **2:**205-10.

Amoutzias, G., D., Robertson, D., L., Van de Peer, Y., and Oliver, S., G. (2008). Choose your partners: dimerization in eukaryotic transcription factors. *Trends Biochem. Sci.* **33**: 220-229.

An, C., and Mou, Z. (2011). Salicylic acid and its function in plant immunity. *J. Integr. Plant Biol.* **53**, 412–428.

Aravind, L., and Koonin, E.,V. (1999). Fold prediction and evolutionary analysis of the POZ domain: Structural and evolutionary relationship with the potassium channel tetramerization domain. *J. Mol. Biol.* **285:** 1353–1361.

Asselbergh, B., De Vleesschauwer, D. and Höfte, M. (2008). Global switches and fine-tuning— ABA modulates plant pathogen defense. *Mol. Plant Microbe Interact*. **21**:709–719.

Bach., M, Schnitzler, J., P, Seitz, H., U. (1993). Elicitor-induced changes in Ca^{2+} influx, K^{+} efflux, and 4-hydroxybenzoic acid synthesis in protoplasts of *Daucus carota* L. *Plant Physiol.* **103**:407–412.

Badis, G., Berger, M., F., Philippakis, A., A., Talukder, S., Gehrke, A., R., Jaeger, S., A., Chan, E., T., Metzler, G., Vedenko, A. and Chen, X. (2009). Diversity and complexity in DNA recognition by transcription factors *Science*, **324**: 1720–1723.

Bagal, U., R., Leebens-Mack, J., H., Lorenz, W., W., Dean, J., F., D. (2012). The *phenylalanine ammonia lyase (PAL)* gene family shows a gymnosperm-specific lineage. *BMC Genomics*, **13**:1-9.

Baktas, Y., and Eulgem, T. (2015). Synthetic plant defense elicitors. *Front. Plant Sci.,* **5**: 1-17.

Balmer, D., Planchamp, C., Mauch-Mani, B. (2013). On the move: induced resistance in monocots. *J. Exp. Bot.* **64**, 1249–1261.

Barabas, O., Ronning, D., R., Guynet, C., Hickman, A., B., Ton-Hoang, B., Chandler, M., and Dyda, F. (2008). Mechanism of IS200/IS605 family DNA transposases: activation and transposon-directed target site selection. *Cell* **132**:208–220.

Bardwell, V., J., and Treisman, R. (1994). The POZ domain: A conserved protein-protein interaction motif. *Genes Dev.* **8:** 1664–1677.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, **129** :823-837.

Barton, J., K., Lolis, E., J. (1985). Simple coordination complexes: Drugs and probe for DNA structure. *A Journal of Critical Discussion of the Current Literature*, **3**: 321-348.

Ben-Gal, I., Shani, A., Gohr, A., Grau, J., Arviv, S., Shmilovici, A., Posch, S. and Grosse, I. (2005). Identification of transcription factor binding sites with variable-order Bayesian networks. *Bioinformatics* **21**: 2657–2666.

Berger, M., F., Badis, G., Gehrke, A., R., Talukder, S., Philippakis, A., A., Pena-Castillo, L., Alleyne, T., M., Mnaimneh, S., Botvinnik, O., B., Chan, E., T. (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**: 1266–1276.

Berger, M., F., Philippakis, A., A., Qureshi, A., M., He, F., S., Estep, P., W. and Bulyk, M., L. (2006). Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat. Biotechnol*. **24**: 1429–1435.

Bi, Y., M., Kenton, P., Mur, L., Darby, R., and Draper, J. (1995). Hydrogen peroxide does not function downstream of salicylic acid in the induction of PR protein expression. *Plant J.* **8**, 235–245.

Bianchi, A., M., Stansel, R., M., Fairall, L., D., Griffith, J., D., Rhodes, D., and de Lange, T. (1999). TRF1 binds a bipartite telomeric site with extreme spatial flexibility. *EMBO J.* **18**: 5735–5744.

Blanco, F., Garreton, V., Frey, N., Dominguez, c., Perez-Acle, T., Van der Straeten, D., Jordana, X. and Holuigue, L. (2005). Identification of NPRI-dependent and independent genes early induced by salicylic acid treatment in Arabidopsis. *Plant Mol. Bio.* **59**: 927-944.

Blasie, C., A., Berg, J., M. (2002). Structure-based thermodynamic analysis of a coupled metal binding-protein folding reaction involving a zinc finger peptide. *Biochemistr*y, **41**:15068-15073.

Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant. Biol*. **60**: 379-406.

Bortz, P., D., S., and Wamhoff, B., R. (2011). Chromatin Immunoprecipitation (ChIP): Revisiting the Efficacy of Sample Preparation, Sonication, Quantification of Sheared DNA, and Analysis via PCR. *PloS one*, **6**: 1-10.

Boyle, P., and Despres C. (2010). Dual-function transcription factors and their entourage: unique and unifying themes governing two pathogenesis-related genes. *Plant. Signal. Behav.* **5** :629-634.

Boyle, P., Le, S., E., Rochon, A., Shearer, H., L., Murmu, J., Chu, J., Y., Fobert, P., R., and Despres, C. (2009). The BTB/POZ domain of the *Arabidopsis* disease resistance protein NPR1 interacts with the repression domain ofTGA2 to negate its function. *Plant Cell* **21**:3700-3713.

Brennan, R., G. and Matthews, B., W.(1989). The HTH DNA-binding motif. *J Biol Chem*, **264**:1903-1906.

Brozek, V., Dias, L., M., Hadano, H., Münks, K., W., Sawada, H., Sirven, C. (2012). Use of Isothiazolecarboxamides to Create Latent Host Defences in a Plant. World Patents, WO 2012084858 A3.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F. (2002). The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proc Natl Acad Sci*, **99**: 9328–9333.

Brunner, F., Rosah, S., Lee, J., Rudd, J., Geiler, J., Kauppinen, C., S., Rasmussen, Scheel, G., Nurnberger, D. (2002). Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora* transglutaminasa. *EMBO J.* **21**: 6681– 6688.

Bulyk, M., L. (2004). Computational prediction of transcription-factor binding site locations. *Genome Bio*, **5** : 1-11.

Burley, S., K. (1996). The TATA box binding protein. *Curr Opin Struct Biol*, **6:**69-75.

Bustos, M., Guiltinan, M., J., Jordano, J., Begum, D., Kalkan, F., A., and Hall, T., C. (1989). Regulation of *beta-glucuronidase* expression in transgenic tobacco plants by an A/T-rich, *cis* acting sequence found upstream of a French bean beta phaseolin gene. *Plant Cell,* **1** :839-853.

Callis, J., Raasch, J., A., Vierstra, R. (1990). Ubiquitin extension proteins of *Arabidopsis thaliana*; structure, localization and expression of their promoters in transgenic tobacco. *J. Biol. Chem.* **265**: 12486-12493.

Canet, J., V., Dobón, A., Roig, A., Tornero, P. (2010). Structure-function analysis of npr1 alleles in *Arabidopsis* reveals a role for its paralogs in the perception of salicylic acid. *Plant Cell Environ.* **33**: 1911–1922.

Cao, H., Bowling, S., A., Gordon, A., S., Dong, X. (1994). Characterization of an *Arabidopsis* mutant that is nonresponsive to inducers of systemic acquired resistance. *Plant Cell,* **6**:1583–1592.

Cao, H., Glazebrook, J., Clarke, J., D., Volko, S., and Dong, X. (1997). The *Arabidopsis* NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell* **88**, 57–63.

Cao, H., Li, X., and Dong, X. (1998). Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. *Proc. Natl. Acad. Sci.,* **95**: 6531-6536.

Carey, M. (1998). The enhanceosome and transcriptional synergy. *Cell*, **92**: 5–8.
Carey, M., F, Peterson, C., L., and Smale, S., T. (2012). Experimental Strategies for the Identification of DNA-Binding Proteins. *Cold Spring Harbor Protoc* **10**: 18-33.

Castresana, C., Garcia-Luque, I., Alonso, E., Malik, V., S., and Cashmore, A., R. (1988). 60th positive and negative elements mediate expression of a photo regulated CAB gene from *Nicotiana plumbaginifolia*. *EM60 J.* **7** :1929-1936.

Chanda, B., Xia, Y., Mandal, M., K., Yu, K., Sekine, K., T., Gao, Q., M. (2011). Glycerol-3-phosphate is a critical mobile inducer of systemic immunity in plants. *Nat. Genet.* **43**, 421–427.

Chaturvedi, R., Krothapalli, K., Makandar, R., Nandi, A., Sparks, A., Roth, M. R. (2008). Plastid ω-3 desaturase-dependent accumulation of a systemic acquired resistance inducing activity in petiole exudates of *Arabidopsis thaliana* is independent of jasmonic acid. *Plant J.* **54**, 106–117.

Chaturvedi, R., Venables, B., Petros, R., A., Nalam, V., Li, M., Wang, X. (2012). An abietane diterpenoid is a potent activator of systemic acquired resistance. *Plant J.* **71**: 161–172.

Chen, J., W., Kuang, J., F., Peng, G., Wan, S., Liu, R., Yang, Z., D., and Deng, H., H. (2012). Molecular cloning and expression analysis of a NPR1 gene from sugarcane. *Pak. J. Bot.,* **44**: 193-200.

Chen, Z., Agnew, J., L., Cohen, J., D., He, P., Shan, L., Sheen, J., Kunkel, B., N. (2007). *Pseudomonas syringae* type III effector AvrRpt2 alters *Arabidopsis thaliana* auxin physiology. *Proc Natl Acad Sci,* **104**: 20131–20136.

Chern, M., S., Fitzgerald, H., A., Canlas, P., E., Navarre D., A., and Ronald, P., C. (2005). Overexpression of a rice NPR1 homolog leads to constitutive activation of defense response and hypersensitivity to light. *Mol. Plant Microbe. Interact*., **18**: 511-520.

Chern, M., S., Fitzgerald, H., A., Yadav, R., C., Canlas, P., E., Dong, X., and Ronald, P., C. (2001). Evidence for a disease-resistance pathway in rice similar to the NPR1 mediated signaling pathway in *Arabidopsis*. *Plant J.,* **27**: 101-113.

Chester, K. S. (1933). The problem of acquired physiological immunity in plants. *Q. Rev. Biol.* **8**: 275–324.

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nurnberger, T., Jones, J., D. (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature.* **448**:497–500.

Chisholm, S.,T., Coaker, G., Day, B., and Staskawicz, B., J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**: 803-814.

Choi, H., W., Lee, D., H., Hwang, B., K. (2009). The pepper calmodulin gene *CaCaM1* is involved in reactive oxygen species and nitric oxide generation required for cell death and the defense response. *Mol. Plant-Microbe Interact.* **22**: 1389 – 1400.

Chong, K., P., Rossall, S., Atong, M. (2009). *In Vitro* Antimicrobial Activity and Fungitoxicity of Syringic Acid, Caffeic Acid and 4-hydroxybenzoic Acid against Ganoderma Boninense. *Journal of Agricultural Science*, **1**: 15-20.

Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P., and de Lange, T. (1995). A human telomeric protein. *Science* **270**: 1663–1667.

Choudhary, D., K., Prakash, A., Johri, B., N. (2007). Induced systemic resistance (ISR) in plants: mechanism of action. *Indian J. Microbiol*., **47**:289–297.

Collas, P., and Dahl, J., A. (2008) Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation. *Frontiers in Bioscience,* **13** :929–943.

Collins, T., Stone, J., R., and Williams, A., J. (2001). All in the family: The BTB/POZ, KRAB, and SCAN domains. *Mol. Cell. Biol.* **21:** 3609–3615.

Conrath, U. (2011). Molecular aspects of defence priming. *Trends Plant Sci.* **16**:524–531.

Conrath, U., Chen, Z., Ricigliano, J., R., and Klessig, D., F. (1995). Two inducers of plant defense responses, 2,6-dichloroisonicotinec acid and salicylic acid, inhibit catalase activity in tobacco. *Proc. Natl. Acad. Sci.* **92** :7143–7147.

Cozzarelli, N., R., and Wang, J., C. (1990). DNA topology and its biological Effects. **20** Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Cress, W., D. and Triezenberg, S., J. (1991). Critical structural elements of the VP16 transcriptional activation domain. *Science* **251** :87–90.

Czarnecka, E., Ingersoll, J., C., and Gurley, W., B. (1992). AT-rich promoter elements of soybean heat shock gene *Gmhsp*lZSE bind two distinct sets of nuclear proteins in vitro. *Plant MOI. Biol.* **19** :985-1000.

Das, P., M., Ramachandran, K., vanWert, J., and Singal, R. (2004). Chromatin Immunopercipitation Assay, *BioTechnique*, **37** :961-969.

Datta, N., and Cashmore, A., R. (1989). Binding of a Pea Nuclear Protein to Promoters of Certain Photo regulated Genes 1s Modulated by Phosphorylation. *The Plant Cell*, **1**: 1069-1077.

Davletova, S., Rizhsky, L., Liang, H., Zhong, S., Oliver, D., J., Coutu, J., Shulaev, V., Schlauch, K., Mittler, R. (2005). Cytosolic *ascorbate peroxidase 1* is a central component of the reactive oxygen gene network of Arabidopsis. *Plant Cell* **17:**268–281.

De la Cruz, F., Frost, L., S., Meyer, R., J., and Zechner, E., L. (2010). Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol. Rev*.**34**:18–40.

De Vos, M., Van Zaanen, W., Koornneef, A., Korzelius, J., P., Dicke, M., Van Loon, L., C., Pieterse, C., M., J. (2006). Herbivore-induced resistance against microbial pathogens in *Arabidopsis*. *Plant Physiology* **142**: 352–363.

De Young, B., J. and Innes, R., W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat. Immunol.* **7**: 1243-1249.

Delaney, T., P., Friedrich, L., and Ryals, J., A. (1995). *Arabidopsis* signal transduction mutant defective in chemically and biologically induced disease resistance. *Proc. Natl. Acad. Sci.* **92**:6602-6606.

Delwel, R., Funabiki, T., Kreider, B., L., Morishita, K. and Ihle, J., N. (1993) Four of the seven zinc fingers of the Evi-1 myeloidtransforming gene are required for sequence-specific binding to GA(C/T) AAGA(T/C)AAGATAA. *Mol. Cell. Biol*., **13**: 4291–4300.

Dempsey, D. A., Vlot, A. C., Wildermuth, M. C., and Klessig, D. F. (2011). Salicylic acid biosynthesis and metabolism. *Arabidopsis* Book **9**, e0156.

Dempsey, D., A., Klessig, D., F. (2012). SOS – too many signals for systemic acquired resistance? *Trends Plant Sci.* **17**:538–545.

Denancé, N., Sánchez-Vallet, A., Goffner, D, and Molina, A. (2013). Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs. *Front. Plant Sci.***4**: 1-12.

Despres, C., Chubak, C., Rochon, A., Clark, R., Bethune, T., Desveaux, D., and Fobert, P., R. (2003). The *Arabidopsis* NPR1 disease resistance protein is a novel cofactor that confers redox regulation of DNA binding activity to the basic domain/leucine zipper transcription factor TGA1. *Plant Cell* **15** :2181-2191.

Després, C., DeLong, C., Glaze, S., Liu, E., Fobert, P., R. (2000). The *Arabidopsis* NPR1/NIM1 protein enhances the DNA binding activity of a

subgroup of the TGA family of bZIP transcription factors. *Plant Cell* **12** :279–290.

Devereux, M., Gresh, N., Piquemal, J., P., Meuwly, M. (2014). A supervised fitting approach to force field parametrization with application to the SIBFA polarizable force field. *J Comput Chem*, **35**:1577-1591.

Dey, B., Thukral, S., Krishnan, S., Chakrobarty, M., Gupta, S., Manghani, C., Rani, V. (2012). DNA–protein interactions: methods for detection and analysis. *Mol Cell Biochem*, **365**:279–299.

Dudev, T., and Lim, C. (2014) Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins. *Chem. Rev.* **114** :538–556.

Durrant, W., E., and Dong, X. (2004). Systemic acquired resistance. *Annu. Rev.Phytopathol,* **42**: 185-290.

Einck, L., Bustin, M. (1985). The intracellular distribution and function of the high mobility group chromosomal proteins. *Exp Cell Res,* **156**:295–310.

Ellenberger, T., E., Brand, C., J., Struhl, K. and Harrison, S., C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein-DNA complex. *Cell* **71**:1223-1237.

Endah, R., G., Beyene, A., Kiggundu, N., van den Berg, U., Schlüter, K., Kunert, and R., Chikwamba. (2008). Elicitor and *Fusarium*-induced expression of NPR1-like genes in banana. *Plant Physiol. Biochem*., **46**:1007-1014.

Eulgem, T., Somssich, I., E. (2007). Networks of WRKY transcription factors in defence signalling. *Curr Opin Plant Biol* **10:** 366–371.

Fan, J., and Doerner, P. (2012). Genetic and molecular basis of non-host disease resistance: complex, yes; silver bullet, no. Curr. Opin. *Plant Biol.* **15**: 400-406.

Fan, W., and Dong, X. (2002). *In vivo* interaction between NPR1 and transcription factor TGA2 leads to salicylic acid–mediated gene activation in *Arabidopsis*. *Plant Cell* **14** :1377–1389.

Fersht, A., R. (1993). The sixth Datta Lecture. Protein folding and stability: The pathway of folding of barnase. *FEBS Lett.* **325** :5–16.

Feuillet, C., Schachermayr, G., Keller, B. (1997). Molecular cloning of a new receptor-like kinase gene encoded at the *Lr10* disease resistance locus of wheat. *Plant J* **11**: 45–52.

Feys, B., J., Moisan, L., J., Newman, M., A., and Parker, J., E. (2001). Direct interaction between the *Arabidopsis* disease resistance signalling proteins, EDS1 and PAD4. *EMBO J.* **20**, 5400–5411.

Feys, B., J., Parker, J., E. (2000). Interplay of signaling pathways in plant disease resistance. *Trends Genet* **16**:449-455.

Finn, R., D, Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S., R., Sonnhammer, E., L., Bateman A. (2006). Pfam: Clans, web tools and services. *Nucleic Acids Res.* **34**: 247-251.

Fitzgerald, H., A., Chern, M., S., Navarre, R., Ronald, P., C. (2004). Overexpression of *(At)*NPR1 in rice leads to a BTH- and environment-induced lesion-mimic/cell death phenotype. *Mol Plant Microbe Interact*, **17:**140-51.

Flors, V., Ton, J., Van Doorn, R., Jakab, G., Garcia-Agustin, P., Mauch-Mani, B. (2008). Interplay between JA, SA and ABA signalling during basal and induced resistance against *Pseudomonas syringae* and *Alternaria brassicicola*. *Plant J.* **54**:81–92.

Foster, A., W., Osman, D., and Robinson, N., J. (2014). Metal Preferences and Metallation. *The Journal of Biological Chemistry,* **289**: 28095–28103.

Fox, K., R., Yan, Y., F. and Gong, B. (1999) DNA sequence recognition by a novel series of minor groove-binding ligands. *Anti-Cancer Drug Des*, **14** :219-230.

Friedrich, L., Lawton, K., Ruess, W., Masner, P., Specher, N., Gut-Rella, M., Meier, B., Dincher, S., Staub, T., Uknes, S. (1996). A benzothiadiazole derivative induces systemic acquired resistance in tobacco. *Plant J* **10**: 61–70.

Friedrich, L., Vernooij, B., Gaffney, T., Morse, A., Ryals, J.(1995). Characterization of tobacco plants expressing a bacterial salicylate hydroxylase gene. *Plant Mol Biol* **29**:959–968.

Fu, Z., Q., Yan, S., Saleh, A., Wang, W., Ruble, J., Oka, N., Dong, X. (2012). NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. *Nature*, **486**: 228–232.

Gaffney, T., Friedrich, L., Vemooij, B., Negrotto, D., Nye, G., UImes, S., Ward, E., Kessmann, H., and Ryals, J. (1993). Requirement of salicylic Acid for the induction of systemic acquired resistance. *Science* **261**, 754-756.

Gajiwala, K., S, Burley, S., K. (2000). Winged helix proteins. *Curr. Opin. Struct. Biol.* **10** 110-116.

Gao, Q., Zhu, S., Kachroo1, P., and Kachroo, A. (2015). Signal regulators of systemic acquired resistance. *Frontiers in Plant Science*, **6**: 1-12.

Garcion, C., Lohmann, A., Lamodière, E., Catinot, J., Buchala, A., Doermann, P. (2008). Characterization and biological function of the

ISOCHORISMATE SYNTHASE2 gene of *Arabidopsis. Plant Physiol*. **147**, 1279–1287.

Garner, M., M. and Revzin, A. (1986). The Use of Gel-Electrophoresis to Detect and Study Nucleic-Acid Protein Interactions. *Trends in biochemical sciences*, **11** :395-396.

Garner, M., M., and Revzin, A. (1981). A Gel-Electrophoresis Method for Quantifying the Binding of Proteins to Specific DNA Regions - Application to Components of the *Escherichia-Coli* Lactose Operon Regulatory System. *Nucleic Acids Res*, **9**, 3047-3060.

Garvie, C., W. and Wolberger, C. (2001). Recognition of specific DNA sequences. *Mol. Cell*, **8**: 937–946.

Giri, J., Vij, S., Dansana, P., K., Tyagi, A., K. (2011). Rice A20/AN1 zinc-finger containing stress-associated proteins (SAP1/11) and a receptor-like cytoplasmic kinase (*Os*RLCK253) interact via A20 zinc-finger and confer abiotic stress tolerance in transgenic *Arabidopsis* plants. *New Phytol.,* **191** :721–732.

Gómez-Gómez, L., and Boller, T. (2002). Flagellin Perception: A Paradigm for Innate Immunity. *Trends in Plant Science* **7**: 251–56.

Gomez-Gomez, L., Bauer, Z., and Boller, T. (2001). Both the extracellular leucine-rich repeat domain and the kinase activity of FSL2 are required for flagellin binding and signaling in *Arabidopsis*. *Plant Cell* **13**: 1155-1163.

Gonzalez-Perez, B., Lucas, M., Cooke, L., A., Vyle, J., S., de la Cruz, F., and Moncalian, G. (2007). Analysis of DNA processing reactions in bacterial conjugation by using suicide oligonucleotides. *EMBO J.* **26**:3847–3857.

Gordan, R., Shen, N., Dror, I., Zhou, T., Horton, J., Rohs, R. and Bulyk, M., L. (2013). Genomic regions flanking E-box binding sites influence DNA binding specificity of bHLH transcription factors through DNA shape. *Cell Reports*, **3**:1093–1104.

Granok, H., Leibovitch, B., A., Shaffer, C., D., and Elgin, S., C. (1995). Chromatin. Ga-ga over GAGA factor. *Curr. Biol.* **5**: 238–241.

Gray, H., B. (2003). Biological inorganic chemistry at the beginning of the 21st century. *Proc Natl Acad Sci* **100**:3563-3568.

Greenberg, J., T., and Yao. (2004). The role and regulation of programmed cell death in plant pathogen interactions. *Cell. Microbiol*. **6**: 201-211.

Gupta, S., K., Rai, A., K., Kanwar, S., S., Sharma, T., R. (2012) Comparative Analysis of Zinc Finger Proteins Involved in Plant Disease Resistance. *PLoS ONE* **7**: 1-15.

Ha, C., M., Jun, J., H., Nam, H., G., and Fletcher, J., C. (2004). BLADE-ON-

PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in Arabidopsis thaliana. *Plant Cell Physiol*. **45** :1361–1370.

Hagman, J., Grossched, R. (1992). An inhibitory carboxyl-terminal domain in Ets-1 and Ets-2 mediates differential binding of ETS family factors to promoter sequences of the mb-1 gene. *Proc Natl Acad Sci* **89** :8889–8893.

Halim, V., Altmann, A., Ellinger, S., Eschen-Lippold, D., Miersch, L., Scheel, O., Rosahl, S. (2009). PAMP-induced defense responses in potato require both salicylic and jasmonic acid. *Plant J.* **57**: 230 – 242.

Hampshire, A., J., Rusling, D., A., Broughton‐Head, V., J. and Fox, K., R. (2007). Footprinting: A method for determining the sequence selectivity, affinity and kinetics of DNA‐binding ligands. *Methods*, **42** :128‐140.

Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C., Stam, M. (2007). Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods*, **3**: 1-16.

Hellman, L., M., and Fried, M., G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein‐nucleic acid interactions. *Nat Protoc*, **2** :1849‐1861.

Henanff, G., L., T., Heitz, P., Mestre, J., Mutterer, B., Walter, and J., L., Chong. (2009). Characterization of Vitis vinifera NPR1 homologs involved in the regulation of Pathogenesis- Related gene expression. *B.M.C. Plant Biol.,* **9**: 1-14.

Hepworth, S., R., Zhang, Y., McKim, S., Li, X., and Haughn, G., W. (2005). BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in *Arabidopsis*. *Plant Cell.* **17** :1434–1448.

Horvath, D., and Chua, N., H. (1996). Identification of an immediate-early salicylic acid inducible tobacco gene and characterization of induction by other compounds. *Plant Mol. Bio.* **31**: 1061-1072.

Horvath, D., M., Huang, D., J. and Chua, N., H. (1998). Four Classes of Salicylate Induced Tobacco Genes. *Mol. Plant Microbe Interact.* **11**: 895-905.

Horváth, E., Pál, M., Szalai, G., Páldi, E., Janda, T. (2007). Exogenous 4-hydroxybenzoic acid and salicylic acid modulate the effect of short-term drought and freezing stress on wheat plants. *Biol. Plant.* **51**: 480–487.

Horwitz, M., S., and Loeb., L., A. (1988). An *E. coli* promoter that regulates transcription by DNA superhelix-induced cruciform extrusion. *Science* **241**: 703–705.

Hu, Y., X., Wang, Y., X., Liu, X., F., Li, J., Y. (2004). *Arabidopsis* RAV1 is down regulated by brassinosteroid and may act as a negative regulator during plant development. *Cell Research* **14**: 8–15.

Huang, J., Gu, M., Lai, Z., Fan, B., Shi, K., Zhou, Y., H. (2010). Functional analysis of the *Arabidopsis* PAL gene family in plant growth, development, and response to environmental stress. *Plant Physiol.* **153**, 1526–1538.

Huang, T., Teng, C., Chang, J., Chuang, H., Ho, C., and Wu, M. (2008). Biosynthetic Mechanism of 2-Acetyl-1-pyrroline and Its Relationship with Δ^1-Pyrroline-5-carboxylic Acid and Methylglyoxal in Aromatic Rice (*Oryza sativa L.*) Callus. *J. Agric. Food Chem.*, ***56*** :7399–7404.

Hwang, R.,Y., Zhong, L., Xu, Y., Johnson, T., Zhang, F., Deisseroth, K., Tracey, W., D. (2007). Nociceptive neurons protect Drosophila larvae from parasitoid wasps. *Curr. Biol.* **17**: 2105-2116.

Iida, A., Kazuoka, T., Torikai, S., Kikuchi, H., Oeda, K. (2000). A zinc finger protein RHL41 mediates the light acclimation response in *Arabidopsis*. *Plant J*, 24: 191–203.

Iriti, M., Faoro, F. (2009). Chitosan as a MAMP, searching for a PRR. *Plant Signal. Behav*. **4**: 66 – 68.

Ishihara, T., Sekine, K. T., Hase, S., Kanayama, Y., Seo, S., Ohashi, Y. (2008). Overexpression of the *Arabidopsis thaliana* EDS5 gene enhances resistance to viruses. *Plant Biol.* **10**, 451–461.

Ito, S., Nakamichi, N., Matsushika, A., Fujimori, T., Yamashino, T., and Mizuno, T. (2005). Molecular dissection of the promoter of the light-induced and circadian-controlled APRR9 gene encoding a clock-associated component of Arabidopsis thaliana. *Bioscience, Biotechnology, and Biochemistry* **69**: 382-390.

Izatt, R., M, Christensen, J., J., Rytting, J., H. (1971). Sites and thermodynamic quantities associated with proton and metal ion interaction with ribonucleic acid, deoxyribonucleic acid, and their constituent bases, nucleosides, and nucleotides. *Chem Rev,* **5**: 439-481.

Jacoby, M., Weisshaar, B., Droge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., Parey, F. (2002). bZIP transcription factors in *Arabidopsis*. *Trends Plant Sci.* **7** :106-111.

Jain, D., Baldi, S., Zabel, A., Straub., T, and Becker, P., B. (2015). Active promoters give rise to false positive 'Phantom Peaks' in ChIP-seq experiments. *Nucleic Acids Research*, **10**:1-10.

Jayaram, B. and Jain, T. (2004). The role of water in protein-DNA recognition. *Annu. Rev. Biophys. Biomol. Struct*. **33**:343–361.

Jelesarov, I., Bosshard, H., R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J Mol Recognit* **12**:3–18.

Jenns, A., and Kuc, J. (1979). Graft transmission of systemic resistance of cucumber to anthracnose induced by *Colletotrichum lagenarium* and tobacco necrosis virus. *Phytopathology*, **69**:753–756.

Jin, J., Shi, J., Liu, B., Liu, Y., Huang, Y., Yu, Y., and Dong, A. (2015). MRG702, a reader protein of H3K4me3 and H3K36me3, is involved in brassinosteroid-regulated growth and flowering time control in rice. *Plant Physiology*, **10**: 1-33.

Johnson, C., Boden, E., and Arias, J. (2003). Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in *Arabidopsis*. *Plant Cell* **15** :1846-1858.

Jones, J., D, G., Dangl, J., L. (2006). The plant immune system. *Nature* **444**: 323– 329.

Jonsen, M., D., Petersen, J., M., Xu, Q., P., Graves, B., J. (1996). Characterization of the cooperative function of inhibitory sequences in Ets-1. *Mol. Cell. Biol.* **16**:2065–2073.

Jun, J., H., Ha, C., M., and Fletcher, J., C. (2010). BLADE-ON-PETIOLE1 coordinates organ determinacy and axial polarity in *Arabidopsis* by directly activating ASYMMETRIC LEAVES2. *Plant Cell.* **22**, 62–76.

Jung, H., W., Tschaplinski, T., J., Wang, L., Glazebrook, J., Greenberg, J., T. (2009). Priming in systemic plant immunity. *Science* **324**, 89–91.

Kagaya, Y., Ohmiya, K., and Hattori, T. (1999). RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.* **27**, 470–478.

Kalodimos, C., G, Boelens, R., Kaptein, R. (2004). Toward an integrated model of protein-DNA recognition as inferred from NMR studies on the Lac repressor system. *Chem. Rev.* **104**:3567–86.

Kalodimos, C., G., Biris, N., Bonvin, A., M., Levandoski, M., M., Guennuegues, M., Boelens, R. (2004). Structure and flexibility adaptation in nonspecific and specific protein–DNA complexes. *Science*, **305** :386–389.

Kant, C., R., Rao, B., J., and Sainis, J., K. (2005). DNA binding and pairing activity of *Os*Dmc1, a recombinase from rice. *Plant Mol. Biol*., **57** :1–11.

Katagiri, F., Lam, E., and Chua, N., H. (1989). Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. *Nature* **340** :727-730.

Kato, M., Hokabe, S., Itakura, S., Minoshima, S., Lyubchenko, Y., L., Gurkov, T., D., Okawara, H., Nagayama, K., and Shimizu, N. (2003). Interarm interaction of DNA cruciform forming at a short inverted repeat sequence. *Biophys. J.* **85**:402-408.

Katsani, K., R., Hajibagheri, M., A., and Verrijzer, C., P. (1999). Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology. *Embo J,* **18** :698-708.

Kesarwani, M., Yoo, J., and Dong, X. (2007). Genetic interactions of TGA transcription factors in the regulation of pathogenesis-related genes and disease resistance in *Arabidopsis*. *Plant Physiol.* **144** :336-346.

Kessman, H., Staub, T., Hofmann, C., Maetzke, T., Herzog, G., Ward, E., Uknes, S., Ryals, J. (1994). Induction of systemic acquired disease resistance in plants by chemicals. *Annu. Rev. Phytopathol.* :439-460.

Khadem, S., and Marles, R., J. (2010). Monocyclic Phenolic Acids; Hydroxy and Polyhydroxybenzoic Acids: Occurrence and Recent Bioactivity Studies. *Molecules*, **15**:7985-8005.

Khan, M., Xu, H., Hepworth, S., R. (2014). BLADE-ON-PETIOLE genes: Setting boundaries in development and defense. *Plant Sciences*, **216**: 157-171.

Khanna, R., Kronmiller, B., Maszle, D., R., Coupland, G., Holm, M., Mizuno, T., Wu, S., H. (2009). The *Arabidopsis* B-box zinc finger family. *Plant Cell* **21**: 3416–3420.

Khorasanizadeh, S. and Rastinejad, F. (2001). Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci.* **26**: 384–390.

Kim, J. and Struhl, K. (1995). Determinants of half-site spacing preferences that distinguish AP-1 and ATF/CREB bZIP domains. *Nucleic Acids Res*. **23**: 2531–2537.

Kinkema, M., Fan, W., and Dong, X. (2000). Nuclear localization of NPR1 is required for activation of PR gene expression. *Plant Cell* **12**, 2339-2350.

Kirchler, T., Briesemeister, S., Singer, M., Schutze, K., Keinath, M., Kohlbacher, O., Vicente-Carbajosa, J., Teige, M., Harter, K., and Chaban, C. (2010). The role of phosphorylatable serine residues in the DNA-binding domain of *Arabidopsis* bZIP transcription factors. *Eur. J. Cell. Bio.* **89**: 175-183.

Kisu, Y., Ono, T., Shinofurutani, N., Suzuki, M., Esaka, M. (1998). Characterization and expression of a new class of zinc finger protein that binds to silencer region of *ascorbate oxidase* gene. *Plant Cell Physiol* **39:** 1054–1064.

Klug, A., and Rhodes, D. (1987). Zinc fingers: a novel protein motif for nucleic acid recognition. *TIBS* **12** :464-469.

Knoth, C., Salus, M. S., Girke, T., and Eulgem, T. (2009). The synthetic elicitor 3,5-dichloroanthranilic acid induces NPR1-dependent and NPR1-independent mechanisms of disease resistance in *Arabidopsis*. *Plant Physiol.* **150**, 333–347.

Köhle, A., Sommer, S., Yazaki, K., Ferrer, A., Boronat, A., Li, S., M., Heide, L. (2002). High-level expression of *chorismate pyruvate-lyase (ubiC)* and *HMG-CoA reductase* in hairy root cultures of *Lithospermum erythrorhizon*. *Plant Cell Physiol* **43**: 894-902.

Koornneef, A., Leon-Reyes, A., Ritsema, T., Verhage, A., Den Otter, F., C, Van Loon, L., C, and Pieterse, C., M., J. (2008). Kinetics of salicylate-mediated suppression of jasmonate signaling reveal a role for redox modulation. *Plant Physiol.* **147**: 1358–1368.

Kuai, X., Mc Leod, B., J., and Despres, C. (2015). Integrating data on the *Arabidopsis* NPR1/NPR3/NPR4 salicylic acid receptors; a differentiating argument. *Front. Plant Sci.* **6**: 1-5.

Kuc, J. (1982). Induced immunity to plant disease. *Bioscience* **32**, 854–860.

Kumar, S., Nei, M., Dudley, J., and Tamura, K. (2008). MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief in Bioinformatics*, **9**: 299-306.

Kunz, W., Schurter, R., and Maetzke, T. (1997). The chemistry of benzothiadiazole plant activators. *Pestic. Sci.* **50**, 275–282.

Lagaert, S., Beliën, T., Volckaert, G. (2009). Plant cell walls: protecting the barrier from degradation by microbial enzymes. *Seminars in Cell & Developmental Biology* **20**: 1064 –1073.

Laity, J., H., Dyson, H., J, Wright, P., E. (2000). DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys2-His2 zinc fingers. *J. Mol. Biol.* **295**:719–27.

Laity, J., H., Lee, B., M., and Wright, P., E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* **11**: 39–46.

Lam, E., and Chua, N., H. (1990). GT-1 binding site confers light responsive expression in transgenic tobacco. *Science,* **248**, 471-474.

Lando, D., Pongratz, I., Poellinger, L., and Whitelaw, M., L. (2000). A redox mechanism controls differential DNA binding activities of hypoxia-inducible factor (HIF) 1alpha and the HIF-like factor. *J. Biol. Chem.* **275**, 4618-4627.

Laporte, L., Thomas, J., George J. (1998). Structural Basis of DNA Recognition and Mechanism of Quadruplex Formation by the β Subunit of the *Oxytricha* Telomere Binding Protein. *Biochemistry*, **37**, 1327-1335.

Lavalley, V., Chaudouët, P., Stambouli, V. (2007). An atomic force microscopy study of DNA hairpin probes monolabelled with gold nanoparticle: Grafting and hybridization on oxide thin films. *Surface Science,* **601:**5424–5432.

Lawso, C., L, Swigon, D., Murakami, K., S, Darst, S., A, Berman, H., M, and Ebright, R., H. (2004). Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* **14**:10–20.

Lawton, K., A., Friedrich, L., Hunt, M., Weymann, K., Delaney, T., Kessmann, H. (1996). Benzothiadiazole induces disease resistance in *Arabidopsis* by activation of the systemic acquired resistance signal transduction pathway. *Plant J.* **10**, 71–82.

Lazaridis, T., Archontis, G., and Karplus, M. (1995). Enthalpic contribution to protein stability: Insights from atom-based calculations and statistical mechanics. *Adv. Protein Chem*. **47** :231–306.

Leavitt, S., Freire, E. (2001). Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr Opin Struct Biol* **11**:560–566.

Lebel, E., Heifetz, P., Thome, L., Dimes, S., Ryals, J., and Ward, E. (1998). Functional analysis of regulatory sequences controlling PR -1 gene expression in *Arabidopsis. Plant J.* **16**, 223-233.

Lee, J., S., Latimer, L., J., P., and Reid, R., S. (1993). A cooperative conformational change in duplex DNA induced by Zn^{2+} and other divalent metal ions. *Biochem. Cell Biol.,* **71**, 162–168.

Lehmeyer, M., Kanofsky, K., Hanko, E., K., R., Ahrendt, S., Wehrs, M., Machens, F., and Hehl, R. (2015). Functional dissection of a strong andspecific microbe-associated molecular pattern-responsive synthetic promoter. *Plant Biotechnology Journal*: 1-11.

Letunic, I., Copley, R.R., Pils, B., Pinkert, S., Schultz, J., and Bork, P. (2006). SMART 5: Domains in the context of genomes and networks. *Nucleic Acids Res.* **34** :257-260.

Leung, C., H., Grill, S., P., Lam, W., Han, Q., B., Sun, H., D., Cheng, Y., C. (2008). Novel mechanism of inhibition of nuclear factor-κB DNA-binding activity by diterpenoids isolated from Isodon rubescens. *Mol Pharmacol*, **68**: 286–97.

Li, Q., Li, N., Hu, X., Li, J., Du, Z., Chen, L., Yin, G., Duan, J., Zhang, H., Zhao, Y. and Wang, J. (2011). Genome-wide mapping of DNA methylation in chicken. *Plos One*, **6** : 1-7.

Li, W., Zhang, J., Wang, J., Wang, W. (2008). Metal-coupled folding of Cys2His2 zinc-finger. *J Am Chem Soc*, **130**:892-900.

Li, X., Peng, H., Schultz, D., C., Lopez-Guisa, J., M., Rauscher III, F., J., and Marmorstein, R. (1999). Structure-function studies of the BTB/POZ

transcriptional repression domain from the promyelocytic leukemia zinc finger oncoprotein. *Cancer Res.* **59:** 5275–5282.

Li, Y., F, Zhu, R., Xu, P. (2005). Activation of the gene promoter of barley β-1,3-glucanase isoenzymes GIII is salicylic acid (SA)-dependent in transgenic rice plants. *J Plant Res* **118**:215-221.

Lickwar, C., R., Mueller, F., Hanlon, S., E., McNally, J., M., and Lieb, J., D. (2012). Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature Letter* **484**: 251-256.

Lim, F., Kraut, N., Framptom, J., Graf, T. (1992). DNA binding by c-Ets-1, but not v-Ets, is repressed by an intramolecular mechanism. *EMBO J.* **11** :643–652.

Lin, W., C., Lu, C., F., Wu, J., W., Cheng, M., L., Lin, Y., M., Yang, N., S., Black, L., Green, S., K., Wang, J., W., and Cheng. C., P. (2004). Transgenic tomato plants expressing the *Arabidopsis* NPR1 gene display enhanced resistance to a spectrum of fungal and bacterial diseases. *Transgenic Res.*, **13**: 567-581.

Lindermayr, C., Sell, S., Müller, B., Leister, D., and Durner, J. (2010). Redox regulation of the NPR1-TGA1 system of Arabidopsis thaliana by nitric oxide. *Plant Cell* **22**:2894-2907.

Liu, H., Boulton, M., I., Davies, J., W. (1997). Maize streak virus coat protein binds single- and double-stranded DNA in vitro. *J. Gen. Virol.,* **78** :1265–1270.

Liu, J., and Wang, X., J. (2006). An integrative analysis of the effects of auxin on jasmonic acid biosynthesis in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* **48**:99–103.

Liu, J., J., Ekramoddoullah, A., K., Piggott, N., Zamani, A. (2005). Molecular cloning of a pathogen/wound-inducible PR-10 promoter from *Pinus monticola* and characterization in transgenic *Arabidopsis* plants. *Planta* **221**: 159–169.

Liu, Y.G., Mitsukawa, N., Oosumi, T., and Whittier, R.F. (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8,** 457–463.

Liu, Z., Li, F., Zhang, B., Li, S., Wu, J., Shi, Y. (2015). Structural Basis of Plant Homeodomain Finger 6 (PHF6) Recognition by the Retinoblastoma Binding Protein 4 (RBBP4) Component of the Nucleosome Remodeling and Deacetylase (NuRD) Complex. *J Biol Chem*. **290**: 6630-6638.

Lopez, M., M., Makhatadze, G., I. (2002). Differential scanning calorimetry. *Methods Mol Biol* **173**:113–119.

Löscher, R., Heide, L. (1994). Biosynthesis of *p*-hydroxybenzoate from *p*-coumarate and *p*-coumaroyl-CoA in cell-free extracts of*Lithospermum erythrorhizon* cell cultures. *Plant Physiol.* **106**: 271–279.

Lu, H., Zhang, C., Albrecht, U., Shimizu, R., Wang, G., and Bowman, K., D. (2013). Overexpression of a citrus NDR1 ortholog increases disease resistance in *Arabidopsis*. *Front. Plant Sci.* **4**:1-10.

Luisi, B., F. (1992). DNA-transcription – zinc standard for economy. *Nature*, **356**:379-380.

Luscombe, N., M, Austin, S., E, Berman, H., M, and Thornton, J., M. (2000). An overview of the structures of protein-DNA complexes *Genome Biol.* **1**(1): 1-37.

Luscombe, N., M., and Thornton, J., M. (2002). Protein–DNA interactions: Amino acid conservation and the effects of mutations on binding specificity. *J. Mol. Biol.* **320**: 991–1009.

Luscombe, N., M., Laskowski, R., A., and Thornton, J., M. (2001). Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* **29**, 2860–2874.

MacDonald, D., Demarre, G., Bouvier, M., Mazel, D., and Gopaul, D., N. (2006). Structural basis for broad DNA specificity in integron recombination. *Nature* **440**:1157–1162.

Mace, H., A., F., Pelham, H., R., B., Travers, A., A. (1983). Association of an S_1 nuclease-sensitive structure with short direct repeats 5′ of *Drosophila* heat shock genes. *Nature*, **304**, 555-557.

Maeda, H., and Dudareva, N. (2012). The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.***63**: 73–105.

Maier, F., Zwicker, S., Huckelhoven, A., Meissner, M., Funk, J., Pfitzner, A., J., Pfitzner, U., M. (2010). Non-expressor of pathogenesis related gene 1 (NPR1) and some NPR1-related proteins are sensitive to salicylic acid. *Mol Plant Pathol* **12**:73–91.

Makandar, R., Essig, J., S., Schapaugh, M., A., Trick, H., N., and Shah, J. (2006). Genetically engineered resistance to *Fusarium* head blight in wheat by expression of *Arabidopsis* NPR1. *Mol Plant Microbe Interact*., **19**: 123-129.

Maldonado, A., M., Doerner, P., Dixon, R., A., Lamb, C., J., Cameron, R., K. (2002). A putative lipid transfer protein involved in systemic acquired resistance signalling in *Arabidopsis*. *Nature* **419**, 399–403.

Malnoy, M., Jin, Q., Borejsza-Wysocka, E., E., He, S., Y., and Aldwinchle, H., S. (2007). Overexpression of the apple *Mp*NPR1 gene confers increased disease resistance in *Malus x domestica. Mol. Plant Microbe. Interact.,* **20**: 1568-

1580.

Manohar, M., Tian, M., Moreau, M., Park, S., W., Choi, H., W., Fei, Z., (2015). Identification of multiple salicylic acid-binding proteins using two high throughput screens. *Front. Plant Sci.***5**: 1-14.

Masai, H., and Arai, K. (1997). Frpo: a novel single-stranded DNA promoter for transcription and for primer RNA synthesis of DNA replication. *Cell* **89**:897–907.

Matías-Hernández, L., Auguilar-Jalamillor, A., E, Marín-González E, Suárez-López, P, Pelaz, S. (2014). RAV genes: regulation of floral induction and beyond. *Ann Bot.* **114**:1459-1470.

Mauch-Mani, B. and Mauch, F. (2005). The role of abscisic acid in plant-pathogen interactions. *Curr. Opin. Plant Biol*. **8**:409–414.

Mauch-Mani, B., and Slusarenko, A., J. (1996). Production of salicylic acid precursors is a major function of phenylalanine ammonia-lyase in the resistance of *Arabidopsis to Peronospora parasitica*. *Plant Cell* **8**, 203–212.

McKenna, N., J., and O'Malley, B., W. (2002). Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell*, **108**: 465‑474.

McKim, S., M., Stenvik, G., E., Butenko, M., A., Kristiansen, W., Cho, S., K., Hepworth, S., R., Aalen, R., B., Haughn, G., W. (2008). The BLADE-ON-PETIOLE genes are essential for abscission zone formation in *Arabidopsis*. *Development* **135**, 1537–1546.

Medzhitov, R. (2001). Toll-like receptors and innate immunity. *Nat. Rev. Immunol*. **1**: 135 – 145.

Meganathan, R. (2001). Ubiquinone biosynthesis in microorganisms. *FEMS Microbiol Lett* **203**: 131-139.

Metraux, J., P., Ahlgoy, P., Staub, T., Speich, J., Steinemann, A., Ryals, J. (1991). Induced systemic resistance in cucumber in response to 2,6-dichloro-isonicotinic acid and pathogens. *Advances in Molecular Genetics of Plant-Microbe Interactions,* **1**.

Metraux, J., P., Signer, H., Ryals, J., Ward, E., Wyss-Benz, M., Gaudin, J. (1990). Increase in salicylic acid at the onset of systemic acquired resistance in cucumber. *Science* **250**, 1004–1006.

Miller, M. (2009). The importance of being flexible: the case of basic region leucine-zipper transcriptional regulators. *Curf. Protein Pept. Sci.* **10**,244-269.

Mirkin, S., M. (2001). DNA Topology: Fundamentals. *Nature Publishing Group*: 1-11.

Mirlohi, A., Brueggeman, R., Drader, T., Nirmala, J., Steffenson, B., J., and Kleinhofs, A. (2008). Allele sequencing of the barley stem rust resistance gene *Rpg1* identifies regions relevant to disease resistance. *Phytopathology* **98**:910–918.

Mishina, T., E., and Zeier, J. (2007). Pathogen-associated molecular pattern recognition rather than development of tissue necrosis contributes to bacterial induction of systemic acquired resistance in Arabidopsis. *Plant J.* **50** :500–513.

Mittler, R. (2006). Abiotic stress, the field environment and stress combination. *Trends Plant Sci* **11**: 15–19.

Mokry, M., Hatzis, P., de Bruijn, E., Koster, J., Versteeg, R., Schuijers, J., van de Wetering, M., Guryev, V., Clevers, H. and Cuppen, E. (2010). Efficient Double Fragmentation ChIP-seq Provides Nucleotide Resolution Protein-DNA Binding Profiles. *Plos One*, **5**: 1-9.

Mou, Z., Fan, W. and Dong, X. (2003). Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell.* **113**: 935-944.

Mukherjee, S., Berger, M., F., Jona, G., Wang, X., S., Muzzey, D., Snyder, M., Young, R., A. and Bulyk, M., L. (2004). Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nature Genetics*, **36** :1331-1339.

Mukhopadhyay, A., Deplancke, B., Walhout, A., J., Tissenbaum, H., A. (2008). Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in *Caenorhabditis elegans*. *Nat Protoc*. **3**:698–709.

Mukhopadhyay, A., Vij, S., Tyagi, A., K. (2004). Overexpression of a zinc-finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. *Proceedings of the National Academy of Sciences*, **101**: 6309–6314.

Müller, S., Kumari, S., Rodriguez, R., Balasubramanian, S. (2010). Small-molecule-mediated G-quadruplex isolation from human cells. *Nat. Chem.* **2** :1095–1098.

Mur, L., A., J, Kenton, P., Lloyd, A., J., Ougham, H., Prats, E. (2008). The hypersensitive response; the centenary is upon us but how much do we know? *J. Exp. Bot*. **59**: 501 – 520.

Mur, L., A., Kenton, P., Atzorn, R., Miersch, O., Wasternack, C. (2006). The outcomes of concentration-specific interactions between salicylate and jasmonate signaling include synergy, antagonism, and oxidative stress leading to cell death. *Plant Physiol* **140**:249-262.

Murchie, A., I., and Lilley, D., M. (1987). The mechanism of cruciform formation in supercoiled DNA: initial opening of central base pairs in salt

dependent extrusion. *Nucleic Acids Res.* **15**:9641–9654.

Nagpal, P., Ellis, C., M., Weber, H., Ploense, S., E. (2005). Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development*, **132** :4107-18.

Najafabadi, H., S, Albu, M., and Hughes, T., R. (2015). Identification of C2H2-ZF binding preferences from ChIP-seq data using RCADE, *Bioinformatics,* **10:** 1–3.

Nakashita, H., Yasuda, M., Nitta, T., Asami, T., Fujioka, S., Arai, Y., Sekimata, K., Takatsuto, S., Yamaguchi, I., Yoshida, S. (2003). Brassinosteroid functions in a broad range of disease resistance in tobacco and rice. *Plant J* **33**: 887–898.

Nandi, A., Welti, R., Shah, J. (2004). The *Arabidopsis thaliana* dihydroxyacetone phosphate reductase gene *SUPPRESSOR OF FATTY ACID DESATURASE DEFICIENCY1* is required for glycerolipid metabolism and for the activation of systemic acquired resistance. *Plant Cell* **16**, 465–477.

Navarro, L., Bari, R., Achard, P., Lison, P., Nemri, A., Harberd, N., P., Jones, J., D., G. (2008). DELLAs control plant immune responses by modulating the balance of jasmonic acid and salicylic acid signaling. *Curr Biol* **18**: 650–655.

Nawrath, C., Métraux, J., P. (1999). Salicylic Acid Induction-Deficient Mutants of Arabidopsis Express PR-2 and PR-5 and Accumulate High Levels of Camalexin After Pathogen Inoculation. *The Plant Cell*, **11**:1393–1404.

Neidle, S. (2008). Principles of Nucleic Acid Structure, 1st edn. (Academic Press, Boston): 38–80.

Nicaise, V., Roux, M., and Zipfel, C. (2009). Recent advances in PAMP-triggered immunity against bacteria: pattern recognition receptors watch over and raise the alarm. *Plant Physiol.* **150**: 1638-1647.

Norberg, M., Holmlund, M., and Nilsson, O. (2005). The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. *Development.* **132** :2203–2213.

Noutoshi, Y., Okazaki, M., Kida, T., Nishina, Y., Morishita, Y., Ogawa, T. (2012). Novel plant immune-priming compounds identified via high-throughput chemical screening target salicylic acid *glucosyltransferases* in *Arabidopsis*. *Plant Cell* **24**, 3795–3804.

Noyes, M., B., Christensen, R., G., Wakabayashi, A., Stormo, G., D., Brodsky, M., H. and Wolfe, S., A. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* **133**: 1277–1289.

Nurnberger, T., and Scheel, D. (2001). Signal transmission in the plant immune

response. *Trends Plant Sci.* **6**: 372-379.

Nurnberger, T., Kemmerling, B. (2006). Receptor protein kinases – pattern recognition receptors in plant immunity. *Trends Plant Sci.* **11**: 519 – 522.

Nutiu, R., Friedman, R., C., Luo, S., Khrebtukova, I., Silva, D., Li, R., Zhang, L., Schroth, G., P. and Burge, C., B. (2011). Direct measurement of DNA affinity landscapes on a high-throughput sequencing instrument. *Nat. Biotechnol*., **29**: 659–664.

Ogawa, M., Kadowak, A., Yamada, T., Kadooka, O. (2011). Applied Development of a Novel Fungicide Isotianil (Stout): Sumitomo Chemical Co., Ltd. *Health & Crop Sciences Research Laboratory, Environmental Health Science Laboratory,*1-16.

Ohlendorf, D., H, Anderson, W., F, Fisher, R., G, Takeda, Y., and Matthews, B., W. (1982). The molecular basis of DNA–protein recognition inferred from the structure of *cro repressor*. *Nature* **298**: 718–713.

Oostendorp, M., Kunz, W., Dietrich, B., and Staub, T. (2001). Induced disease resistance in plants by chemicals. *Eur. J. Plant Pathol.* **107**, 19–28.

Pace, C., N., Shirley, B., A., McNutt, M., and Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. *FASEB J.* **10** :75–83.

Pallas, J., A., Paiva, N., L., Lamb, C., and Dixon, R., A. (1996). Tobacco plants epigenetically suppressed in phenylalanine ammonia-lyase expression do not develop systemic acquired resistance in response to infection by tobacco mosaic virus. *Plant J.* **10**, 281–293.

Panayotatos, N., and Wells, R., D. (1981). Cruciform structures in supercoiled DNA. *Nature*, **289**:466–470.

Pape, S., Thurow, C., Gatz, C. (2010). The *Arabidopsis* PR-1 promoter contains multiple integration sites for the coactivator NPR1 and the repressor SNI1. *Plant Physiology*, **154**: 1805-1818.

Parekh, B., Hatfield, G., W. (1996). Transcriptional activation by protein-induced DNA bending: evidence for a DNA structural transmission model. *Proceedings of the National Academy of Science*, **93**: 1173-1177.

Park, P.J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nature reviews. Genetics*, **10** :669-680.

Park, S., W., Kaimoyo, E., Kumar, D., Mosher, S., and Klessig, D., F. (2007). Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. *Science* **318**, 113–116.

Pavletich, N., P, and Pabo, C., O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268–DNA complex at 2.1A ˚. *Science* **252**: 809–817.

Pearson, W., R. (2013). An Introduction to Sequence Similarity ("Homology") Searching. *Curr Protoc Bioinformatics*. **03**: 1-9.

Perez-Burgos, L., Peters, A., H., Opravil, S., Kauer, M., Mechtler, K., Jenuwein, T. (2004). Generation and characterization of methyllysine histone antibodies. *Methods Enzymol* **376**: 234–254.

Pessler, F., and Hernandez, N. (2003). Flexible DNA binding of the BTB/POZ-domain protein FBI-1. *J. Biol. Chem.* **278:** 29327–29335.

Peterson, B., R., Sun, L., J., and Verdine, G., L. (1996). A critical arginine residue mediates cooperativity in the contact interface between NFAT and AP-1. *Proc. Natl. Acad. Sci.* **93** :13671–13676.

Philippakis, A., A., Qureshi, A., M., Berger, M., F. and Bulyk, M., L. (2008). Design of compact, universal DNA microarrays for protein binding microarray experiments. *J Comput Biol*, **15** :655-665.

Pieterse, C., M., J., Leon-Reyes, A., van der Ent S, van Wees, S., C., M. (2009) Networking by small-molecule hormones in plant immunity. *Nature Chem. Biol.* **5**: 308 – 316.

Pillai, S., and Chellappan, S., P. (2009). ChIP on chip assays: genome-wide analysis of transcription factor binding and histone modifications. *Methods Mol Biol*, **523** :341- 366.

Pilotti, M., A., Brunetti, A., Gallelli, and S., Loreti. (2008). NPR1-like genes from cDNA of rosaceous trees: cloning strategy and genetic variation. *Tree Genet. Genomes*, **4**: 49-63.

Platt, J., R. (1955). Possible separation of intertwined nucleic acid chains by transfer-twist. *Proc. Natl. Acad. Sci.* **41**:181–183.

Pozo, M.J., Van der Ent, S., Van Loon, L., C., and Pieterse, C., M., J. (2008). Transcription factor MYC2 is involved in priming for enhanced defense during rhizobacteria-induced systemic resistance in Arabidopsis thaliana. *New Phytol*. **180**: 511–523.

Priest, H., D., Filichkin, S., A., and Mockler, T., C. (2009). cis-Regulatory elements in plant cell signaling. *Curr Opin Plant Biol*, **12** :643-649.

Putterill, J., Robson, F., Lee, K., Simon, R., Coupland, G. (1995). The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* **80**: 847–857.

Quilis, J., Peñas, G., Messeguer, J., Brugidou, C., San Segundo, B. (2008). The *Arabidopsis* AtNPR1 inversely modulated defence responses against fungal, bacterial or viral pathogens while conferring hypersensitivity to abiotic stresses in transgenic rice. *Molec Plant-Microbe Interact*, **21:**1215-1231.

Racki, L., R. (2003). NPR2: a novel *Arabidopsis* defence response gene. *American Society of Plant Biologists*. Plant pathogen/ Symbiont interaction.

Raha, D., Hong, M. and Snyder, M. (2010). ChIP-Seq: a method for global identification of regulatory elements in the genome. *Current protocols in molecular biology*, Chapter 21, Unit 21.19.1-14.

Rasmussen, M., W., Roux, M., Petersen, M., Mundy, J. (2012). MAP kinase cascades in *Arabidopsis* innate immunity. *Frontiers in Plant Sci*. **3**: 169.

Raumann, B., E, Rould, M., A, Pabo, C., O, and Sauer, R., T. (1994). DNA recognition by β-sheets in the Arc repressor-operator crystal structure. *Nature*, **367**:754-757.

Rizhsky, L., Liang, H., Shuman, J., Shulaev, V., Davletova, S., Mittler, R. (2004). When defence pathways collide: the response of *Arabidopsis* to a combination of drought and heat stress. *Plant Physiol* **134:** 1683–1696.

Roastas, M., and Turlings, T., C., J. (2008). Induction of systemic acquired resistance in *Zea mays*also enhances the plant's attractiveness to parasitoids. *Biol. Control* **46**: 178–186.

Robbe, K., and Bonnefoy, E. (1998). Non-B-DNA structures on the interferon-b promoter? *Biochimie* **80**:665–671.

Robert-Seilaniantz, A., Navarro, L., Bari, R., Jones, J., D., G. (2007). Pathological hormone imbalances. *Curr Opin Plant Biol* **10**: 372–379.

Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y.J., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O.L., He, A., Marra, M., Snyder, M. and Jones, S. (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods*, **4**, 651-657.

Rochon, A., Boyle, P., Wignes, T., Fobert, P. R., and Despres, C. (2006). The coactivator function of *Arabidopsis* NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. *Plant Cell* **18**, 3670–3685.

Rodriguez, F., I., Esch, J., J., Hall, A., E., Binder, B., M., Schaller, G., E., Bleecker, A., B. (1999). A copper cofactor for the ethylene receptor ETR1 from *Arabidopsis*. *Science* **283** :996–998.

Rohs, R., Jin, X., West, S., M., Joshi, R., Honig, B. and Mann, R., S. (2010). Origins of Specificity in Protein-DNA Recognition. *Annu. Rev. Biochem.* **79**:233–69.

Rohs, R., West, S., M., Sosinsky, A., Liu, P., Mann, R., S. and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* **461**:1248–1253.

Rojas, M., R., Noueiry, A., O., Lucas, W., J., Gilbertson, R., L. (1998). Bean dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. *Cell*, **95** :105–113.

Ross, A. F. (1966). "Systemic effects of local lesion formation," in Viruses of Plants, eds A. B. R. Beemster and J. Dijkstra (Amsterdam: NorthHolland Publishing): 127–150.

Ross, P., Weihofen, W., Siu, F., Xie, A., Katakia, H., Wright, S., K, Hunt, I., Brown, R., K., Freire, E. (2015). Isothermal chemical denaturation to determine binding affinity of small molecules to G-protein coupled receptors. *Analytical Biochemistry* **473**: 41–45.

Rushton, P., J., Torres, J., T., Parniske, M., Wernert, P., Hahlbrock, K., Somssich, I., E. (1996). Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J*, **15**:5690–5700.

Rutledge, R., G., Stewart, D. (2008). A kinetic-based sigmoidal model for the polymerase chain reaction and its application to high-capacity absolute quantitative real-time PCR. *BMC Biotechnol*.**8**:1-28.

Ryan, C., A., Huffker, A., Yamaguchi, Y. (2007). New insights into innate immunity in *Arabidopsis*. *Cell Microbiol*. **9**: 1902 – 1908.

Sakamoto, A., Ueda, M., Morikawa, H. (2002). *Arabidopsis* glutathione-dependent formaldehyde dehydrogenase is an *S-nitrosoglutathione reductase*. *FEBS Lett* **515** :20–24.

Sakamoto, H., Maruyama, K., Sakuma, Y., Meshi, T., Iwabuchi, M., Shinozaki, K., Yamaguchi-Shinozaki, K. (2004). *Arabidopsis* Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. *Plant Physiology*, **136** :2734-2746.

Sakharov, D., V., Lim, C. (2005). Zn protein simulations including charge transfer and local polarization effects. *J Am Chem Soc*,**127**:4921-4929.

Saleh, O., Issman, N., Seumel, G., I., Stav, R., Samach, A., Reski, R., Frank, W., Arazi, T. (2011). MicroRNA534a control of *BLADE-ON-PETIOLE 1* and *2* mediates juvenile-to-adult gametophyte transition in *Physcomitrella patens*. *Plant J.* **65**: 661–674.

Sandhu, D., Tasma, I., M., Frasch, R., and Bhattacharyya, M., K. (2009). Systemic acquired resistance in soybean is regulated by two proteins, orthologous to *Arabidopsis* NPR1. *BMC Plant Biol*. **9**:1-14.

Schnitzler, J., P, Madlung, J, Rose, A., Seitz, H., U. (1992). Biosynthesis of *p*-hydroxybenzoic acid in elicitor-treated carrot cell cultures. *Planta*, **188:594**–600.

Schnitzler, J., P., Seitz, H., U. (1989). Rapid responses of cultured carrot cells and protoplasts to an elicitor from the cell wall of *Pythium aphanidermatum* (Edson) Fitzp. *Z Naturforsch*, **44**:1020–1028.

Schultz, S., C, Shields, G., C and Steitz, T., A. (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90◦. *Science* **253**:1001–7.

Schurter, R., Kunz, W., and Nyfeler, R. (1993). Process and a composition for immunizing plants against diseases. US Patent, US5190928A.

Segonzac, C., and Zipfel, C. (2011). Activation of plant pattern-recognition receptors by bacteria. *Curr. Opin. Microbiol.* **14**: 54-61.

Serrano, M., Wang, B., Aryal, B., Garcion, C., Abou-Mansour, E., Heck, S. (2013). Export of salicylic acid from the chloroplast requires the MATElike transporter EDS5. *Plant Physiol.* **162**, 1815–1821.

Shah, J., and Zeier, J. (2013). Long-distance communication and signal amplification in systemic acquired resistance. *Front Plant Sci,* **4**: 1-16.

Shamsi, M., H, and Kraatz, H., B. (2013). Interactions of Metal Ions with DNA and Some Applications. *J Inorg Organomet Polym*, **23**:4–23.

Shearer, H., L., Cheng, Y., T., Wang, L., Liu, J., Boyle, P., Després, C., Zhang, Y., Li, X., and Fobert, P., R. (2012). Arabidopsis Clade I TGA Transcription Factors Regulate Plant Defences in an NPR1-Independent Fashion. *Mol. Plant Microbe Interact.* **25**:1459-1468.

Shi, Z., Maximova, S., Liu, Y., Verica, J., Guiltinan, M., J. (2013). The salicylic acid receptor NPR3 is a negative regulator of the transcriptional defence response during early flower development in *Arabidopsis*. *Mol. Plant*, **6**: 802–816.

Shimono, M., Sugano, S., Nakayama, A., Jiang, C., J., Ono, K., Toki, S. (2007). Rice WRKY45 plays a crucial role in benzothiadiazole-inducible blast resistance. *Plant Cell* **19**, 2064–2076.

Sierk, M., L., Zhao, Q. and Rastinejad, F. (2001). DNA deformability as a recognition feature in the reverb response element. *Biochemistry* **40**:12833–43.

Siggers, T. and Gordon, R. (2013). Protein-DNA binding: complexities and multi-protein codes. *Nucleic Acid Research* **10** (1093): 1-13.

Silverman, P., Seskar, M., Kanter, D., Schweizer, P., Metraux, J.P., and Raskin, I. (1995). Salicylic acid in rice (biosynthesis, conjugation, and possible role). *Plant Physiol.* **108**, 633–639.

Sinden, R., R. (1994). DNA Structure and Function. Academic Press, USA.

Singh, V., Roy, S., Giri, M. K., Chaturvedi, R., Chowdhury, Z., Shah, J. (2013). *Arabidopsis thaliana* FLOWERING LOCUS D is required for systemic acquired resistance. *Mol. Plant Microbe Interact.* **26**, 1079–1088.

Sircar, D., Mitra, A. (2009). Accumulation of *p*-hydroxybenzoic acid in hairy roots of*Daucus carota* 2: Confirming biosynthetic steps through feeding of inhibitors and precursors. *Journal of Plant Physiology,* **166**: 1370–1380.

Sissoeff, I., Grisvard, J. and Guille, E. (1976). Studies of metal ions - DNA interactions: specific behaviour of reiterative DNA sequences. *Prog. Biophys. Mol. Biol*., **31** :165-199.

Smith-Becker, J., Marois, E., Huguet, E., J., Midland, S., L., Sims, J., J., Keen, N., T. (1998). Accumulation of salicylic acid and 4-hydroxybenzoic acid in phloem fluids of cucumber during systemic acquired resistance is preceded by a transient increase in phenylalanine ammonia-lyase activity in petioles and stems. *Plant Physiol*, **116**:231 – 238.

Solomon, M., J., Larsen, P., L., and Varshavsky, V. (1988). Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell,* 53: 937- 947.

Somers, W., S, and Phillips, S., E., V. (1992). Crystal structure of the Met repressor- operator complex at 2.8 A^{o} resolution reveals DNA recognition by beta strands. *Nature*, **359**:387-391.

Sponer, J., Burda, J., V., Sabat, M., Leszczynski, J., and Hobza, P. (1998). Interaction between the Guanine-Cytosine Watson-Crick DNA base pair and hydrated gropu IIa (Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}) and group IIb (Zn^{2+}, Cd^{2+}, Hg^{2+}) metal cations. *J Phys. Chem. A,* **102**: 5951–5957.

Steitz, T., A. (1990). Structural studies of protein–nucleic acid interaction: the sources of sequence-specific binding. *Quarterly Review of Biophysics* **23**: 205–280.

Stellberger, T., Häuser, R., Baiker, A., Pothineni, V., R., Haas, J., Uetz, P. (2010). Improving the yeast two-hybrid system with permutated fusions proteins: the Varicella Zoster Virus interactome. *Proteome Science*, **8**:1-9.

Strauch, A., M. (2001). Protein-DNA Complexes. *Nature publishing group*.

Strawn, M., A., Marr, S., K., Inoue, K., Inada, N., Zubieta, C., and Wildermuth, M., C. (2007). *Arabidopsis isochorismate* synthase functional in pathogen induced salicylate biosynthesis exhibits properties consistent with a role in diverse stress responses. *J. Biol. Chem.* **282,** 5919–5933.

Sugano, S., Jiang, C., J., Miyazawa, S., Masumoto, C. (2010). Role of *Os*NPR1 in rice defence program as revealed by genome-wide expression analysis. *Plant Mol Biol*,**74**:549-62.

Sullivan, K., M., and Lilley, D., M. (1987). Influence of cation size and charge on the extrusion of a salt-dependent cruciform. *J. Mol. Biol*. **193**:397–404.

Suzuki, T. (1994). Abalone myoglobins evolved from *indoleamine dioxygenase*: the cDNA-derived amino acid sequence of myoglobin from *Nordotis madaka*. *J Protein Chem.* **1**:9–13.

Szarka A., Lőrincz, T. (2014). The role of ascorbate in protein folding. *Protoplasma.* **251**: 489-497.

Tada, Y., Spoel, S., H., Pajerowska-Mukhtar, K., Mou, Z., Song, J., Wang, C. (2008). Plant immunity requires conformational charges of NPR1 via *S*-nitrosylation and thioredoxins. *Science* **321** :952–956.

Tateno, M., Yamasaki, K., Amano, N., Kakinuma, J. and Koike, H. (1997). DNA recognition by beta-sheets *Biopolymers* **44**:335–59.

Teytelmana, L., Thurtlec, D., M., Rinec, J., R, and van Oudenaardena, A. (2013). Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *PNAS* **110**: 18602–18607.

Thompson, J., D. Gibson, T., J. Plewniak, F., Jeanmougin, F., and Higgins, D., G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*., **25** :4876–4882.

Toledano, M., B., Leonard, W., J. (1991). Modulation of transcription factor NF-kappa B binding activity by oxidation-reduction in vitro. *Proc. Nat. Acad. Sci.* **88**: 4328–4332.

Truman, W., and Glazebrook, J. (2012). Co-expression analysis identifies putative targets for CBP60g and SARD1 regulation. *BMC Plant Biol.* **12**: 1-17.

Tsai, N., P., Wilkerson, J., R., Guo, W., Maksimova, M., A., DeMartino, G., N., Cowan, C., W. (2012). Multiple autis-linked genes mediate synapse elimination via proteasomal degradation of a synaptic scaffold PSD-95. *Cell* **151** :1581–1594.

Tsuda, K., Sato, M., Glazebrook, J., Cohen, J., D., and Katagiri, F. (2008). Interplay between MAMP-triggered and SA-mediated defense responses. *Plant J.* **53**: 763-775.

Tuzun, S., and Kuc, J. (1985). Movement of a factor in tobacco infected with *Peranospora tabacina* Adam which systemically protects against blue mold. *Physiol. Plant. Pathol.* **26** :321–330.

Tzin, V., and Galili, G. (2010). The biosynthetic pathways for shikimate and aromatic amino acids in *Arabidopsis thaliana*. In The Arabidopsis Book, K. Torii, ed. (Rockville, MD: The American Society of Plant Biologists), p. e0132.

Uknes, S., Mauch-Mani, B., Moyer, M., Potter, S., Williams, S., Dincher, S. (1992). Acquired resistance in *Arabidopsis*. *Plant Cell* **4**, 645–656.

Umemura, Y., Ishiduka, T., Yamamoto, R., Esaka, M. (2004). The Dof domain, a zinc finger DNA-binding domain conserved only in higher plants, truly functions as a Cys2/Cys2 Zn finger domain. *Plant J: Cell Mol Biol.* **37** :741–9.

Uquillas, C., Letelier, I., Blanco, F., Jordana, X. and Holuigue, L. (2004). NPR1- Independent activation of immediate-early salicylic acid-responsive genes in *Arabidopsis*. *Mol. Plant Microbe Interact.* **17**: 34-42.

Val, M., E., Bouvier, M., Campos, J., Sherratt, D., Cornet, F., Mazel, D., and Barre, F., X. (2005). The single-stranded genome of phage CTX is the form used for integration into the genome of Vibrio cholerae. *Mol. Cell* **19**:559–566.

Van Loon, L., C, and Glick, B., R. (2004) Increased plant fitness by rhizobacteria. In Molecular ecotoxicology of plants (*Sandermann H eds*.) *Springer-Verlag, Berlin: Heidelberg* :177–205.

Van Loon, L., C., Rep, M., and Pieterse, C., M., J. (2006). Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol*. **44**: 135–162.

Van Wees, S., C., M., Van der Ent, S., and Pieterse, C., M., J. (2008). Plant immune responses triggered by beneficial microbes. *Curr. Opin. Plant Biol.* **11**: 443–448.

Verberne, M., C., Hoekstra, J., Bol, J., F., Linthorst, H., J., M. (2003). Signaling of systemic acquired resistance in tobacco depends on ethylene perception. *Plant J*.**35**: 27–32.

Vernooij, B., Friedrich, L., Morse, A., Reist, R., Kolditz-Jawhar, R., Ward, E. (1994). Salicylic acid is not the translocated signal responsible for inducing systemic acquired resistance. *Plant Cell* **6** :959–965.

Vij, S., Tyagi, A., K. (2006). Genome-wide analysis of the stress associated protein (SAP) gene family containing A20 ⁄ AN1 zinc-finger(s) in rice and their phylogenetic relationship with *Arabidopsis. Molecular Genetics and Genomics* **276**: 565–575.

Vinson, C., Myakishev, M., Acharya, A., Mir, A., A., Moll, J., R. and Bonovich, M. (2002). Classification of human B-ZIP proteins based on dimerization properties. *Mol. Cell. Biol*. **22**: 6321–6335.

Vlot, A., C., Dempsey, D., A., Klessig, D., F. (2009). Salicylic Acid, a multifaceted hormone to combat disease. *Annu. Rev. Phytopathol.* **47**: 177–206.

Vogel, J., T., Zarka, D., G., Van Buskirk, H., A., Fowler, S., G., Thomashow, M., F. (2005). Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. *Plant J* **41:** 195–211.

Wadkins, R., M. (2000). Targeting DNA Secondary Structures. *Current Medicinal Chemistry*, **7**: 1-15.

Waga, S., Mizuno, S., Yoshida, M. (1990). Chromosomal protein HMG1 removes the transcriptional block caused by the cruciform in supercoiled DNA. *J Biol Chem.* **265** :19424–19428.

Waldron, K., J., Rutherford, J., C., Ford, D., and Robinson, N., J. (2009). Metalloproteins and metal sensing. *Nature* **460**, 823–830.

Wang, C., El-Shetehy, M., Shine, M., B., Yu, K., Navarre, D.,Wendehenne, D. (2014). Free radicals mediate systemic acquired resistance. *Cell Rep.* **7**, 348–355.

Wang, D., Pajerowska-Mukhtar, K., Hendrickson Culler, A., and Dong, X. (2007). Salicylic acid inhibits pathogen growth in plants through repression of the auxin signaling pathway. *Curr. Biol.* **17**, 1784–1790.

Wang, D., Weaver, N., D., Kesarwani, M., and Dong, X. (2005). Induction of protein secretory pathway is required for systemic acquired resistance. *Science* **308** :1036–1040.

Wang, J., Shine, M., B., Gao, Q., M., Navarre, D., Jiang, W., Liu, C. (2014). Enhanced Disease Susceptibility mediates pathogen resistance and virulence function of a bacterial effector in soybean. *Plant Physiol*. **165**, 1269–1284.

Wang, Y., Brindle, I., D. (2011). Ultra-trace determination of vanadium in lake sediments: a performance comparison using O2, N2O, and NH3 as reaction gases in ICP-DRC-MS. *J. Anal. At. Spectrom.,* **26**: 1514–1520.

Wang, Z., X., Yano, M., Yamanouchi, U., Iwamoto, M., Monna, L. (1999). The *Pib* gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. *Plant J*, **19**: 55–64.

Ward, E., R., Uknes, S., J., Williams, S., C., Dincher, S., S., Wiederhold, D., L., Alexander, D. C. (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell,* **3**, 1085–1094.

Wells, R., D., Collier, D., A., Hanvey, J., C., Shimizu, M., Wohlrab, F. (1988). The chemistry and biology of unusual DNA structures adopted by oligopurine oligopyrimidine sequences. *FASEB J* **2**: 2939-2949.

White, J., H., and Bauer, W., R. (1987). Superhelical DNA with local substructures: a generalization of the topological constraint in terms of the intersection number and the ladder-like correspondence surface. *J. Mol. Biol*.**195**:205–213.

White, R., F. (1979). Acetylsalicylic acid (aspirin) induces resistance to tobacco mosaic virus in tobacco. *Virology,* **99** :410–412.

Widhalm, J., R., and Dudareva, N. (2015). A Familiar Ring to It: Biosynthesis of Plant Benzoic Acids. *Molecular Plant*, **8**: 83–97.

Wildermuth, M., C. (2006). Variations on a theme: synthesis and modification of plant benzoic acids. *Curr. Opin. Plant Biol.* **9**:288–296.

Wildermuth, M., C., Dewdney, J., Wu, G., and Ausubel, F., M. (2001). *Isochorismate synthase* is required to synthesize salicylic acid for plant defense. *Nature* **414**, 562–565.

Wilkins, R. C., and Lis, J., T. (1997). Dynamics of potentiation and activation: GAGA factor and its role in heat shock gene regulation. *Nucleic Acids Res.* **25**: 3963–3968.

Williamson, J., R. (1994). G-quartet structures in telomeric DNA. *Annu Rev Biophys Biomol Struct.* **23**:703–730.

Wittung-Stafshede, P. (2002). Role of cofactors in protein folding. *Acc Chem Res*, **35**:201-208.

Wolberger, C. (1999). Multi-protein-DNA complexes in transcriptional regulation. *Annu. Rev. Biophys. Biomol. Struct.*, **28**: 29–56.

Wolfe, S., A., Nekludova, L. and Pabo, C., O. (1999). DNA recognition by Cys2His2 zinc finger proteins. *Annual Review of Biophysical Biochemical Structures* **3**: 183–212.

Wray, G., A., Hahn, M.,W., Abouheif, E., Balhoff, J., P., Pizer, M., Rockman, M.,V. and Romano, L., A. (2003). The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol,* **20** :1377-1419.

Wu, R., Lu, Z., Cao, Z., Zhang, Y. (2011). A transferable nonbonded pairwise force field to model zinc interactions in metalloproteins. *J Chem Theory Comput*, **7**:433-443.

Wu, Y., Zhang, D., Chu, J., Y., Boyle, P., Wang, Y., Brindle, I. D., DeLuca, V., and Despres, C. (2012). The *Arabidopsis* NPR1 protein is a receptor for the plant defence hormone salicylic acid. *Cell Rep.* **1**:639–647.

Xiao, Z., Wedd, A., G. (2010). The challenges of determining metal-protein affinities. *Nat. Prod. Rep*., **27** :768–789.

Xiong, Y., Sundaralingam, M. (2001). Protein–Nucleic Acid Interaction: Major Groove Recognition Determinants *Nature publishing group.*

Yasuda, M., Ishikawa, A., Jikumaru, Y., Seki, M., Umezawa, T., Asami, T., Maruyama-Nakashita, A., Kudo, T., Shinozaki, K., Yoshida, S. (2008). Antagonistic interaction between systemic acquired resistance and the abscisic acid-mediated abiotic stress response in *Arabidopsis*. *Plant Cell*. **20** :1678-1692.

Yoshida, T., Itou, A., Yamamoto, R., Tobino, T., Murakawa, H., and Toda, K. (2013). Determination of isotianil in brown rice and soil using supercritical fluid extraction and gas chromatography/mass spectrometry' *Analytical Sciences*, **29**: 919-922.

Yu, J., Gao, J., Wang, X. Y., Wei, Q., Yang, L. F., Qiu, K. (2010). The pathway and regulation of salicylic acid biosynthesis in probenazole-treated *Arabidopsis*. *J. Plant Biol.* **53**, 417–424.

Yuan, Y., Zhong, S., Li, Q., Zhu, Z., Lou, Y., Wang, L., Wang, J., Wang, M., Li, Q., Yang, D., and He., Z. (2007). Functional analysis of rice NPR1-like genes reveals that *Os*NPR1/NH1 is the rice orthologue conferring disease resistance with enhanced herbivore susceptibility. *Plant Biotechnol. J.*, **5**: 313-324.

Zhang, Y., Cheng, Y., T., Qu, N., Zhao, Q., Bi, D., Li, X. (2006). Negative regulation of defence responses in *Arabidopsis* by two NPR1 paralogs. *Plant J.* **48** :647–656.

Zhang, Y., Fan, W., Kinkema, M., Li, X. and Dong, X. (1999). Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-l gene. *Proc. Natl. Acad. Sci.* **96**: 6523-28.

Zhang, Y., Tessaro, M.J., Lassner, M. and Li, X. (2003). Knockout analysis of Arabidopsis transcription factor TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell.* **15**: 2647-2653.

sZhang, Y., X., Wang, C., Cheng, Q., Gao, J., Liu, and X., Guo. (2008). Molecular cloning and characterization of *Gh*NPR1, a gene implicated in pathogen responses from cotton (*Gossypium hirsutum L.). Biosci. Rep*., **28**: 7-14.

Zhang, Y., Xu, S., Ding, P., Wang, D., Cheng, Y. T., He, J. (2010). Control of salicylic acid synthesis and systemic acquired resistance by two members of a plant-specific family of transcription factors. *Proc. Natl. Acad. Sci.* **107**, 18220–18225.

Zhao, J., Bacolla, A., Wang, G., and Vasquez, K., M. (2010). Non-B DNA structure-induced genetic instability and evolution. Cell. *Mol. Life Sci.*

Zhao, J., T., X. Huang, Y., P., Chen, Y., E., Chen and X., L., Huang. (2009). Molecular cloning and characterization of an ortholog of NPR1 Gene from Dongguan Dajiao *(Musa spp.ABB). Plant Mol. Biol. Rep.,* **27**: 243-249.

Zhou, J., M., Trifa, Y., Silva, H., Pontier, D., Lam, E., Shah, J., and Klessig, D., F. (2000). NPRI differentially interacts with members of the TGA1OBF family of transcription factors that bind an element of the PR-1 gene required for induction by salicylic acid. *Mol. Plant Microbe Interact.* **13** :191-202.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J., D., J., Boller, T,

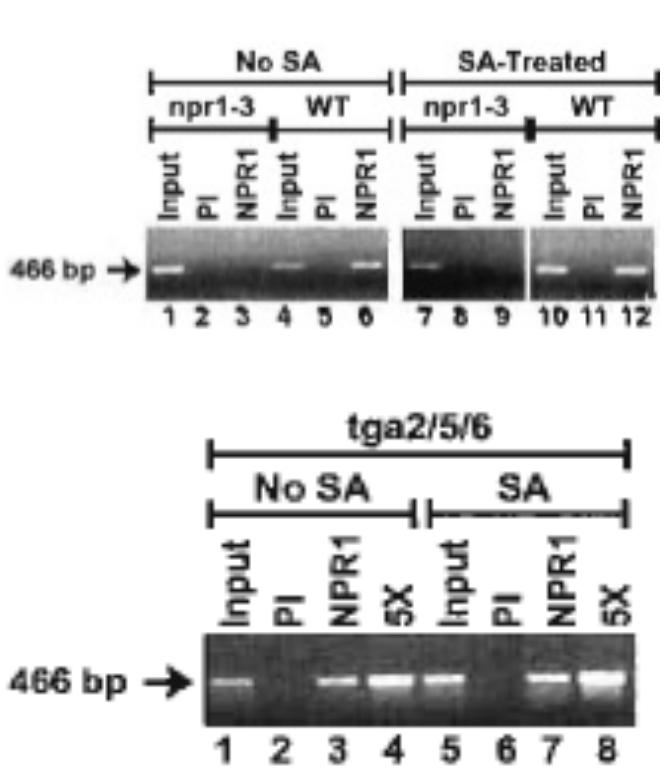
and Felix, G. (2006). Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts *Agrobacterium*-Mediated Transformation. *Cell* **125**: 749–60.

Zschiesche, W., Barth, O., Daniel, K., Böhme, S., Rausche, J., and Humbeck, K. (2015). The zinc-binding nuclear protein HIPP3 acts as an upstream regulator of the salicylate-dependent plant immunity pathway and of flowering time in *Arabidopsis* thaliana. *New Physiologist* **207**: 1084-1096.

Zubay, G. and Doty, P. (1959). Isolation and properties of deoxynucleoprotein particles containing single nucleic acid molecules. *J. Mol. Biol.* **1**:1.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31** :3406–3415.

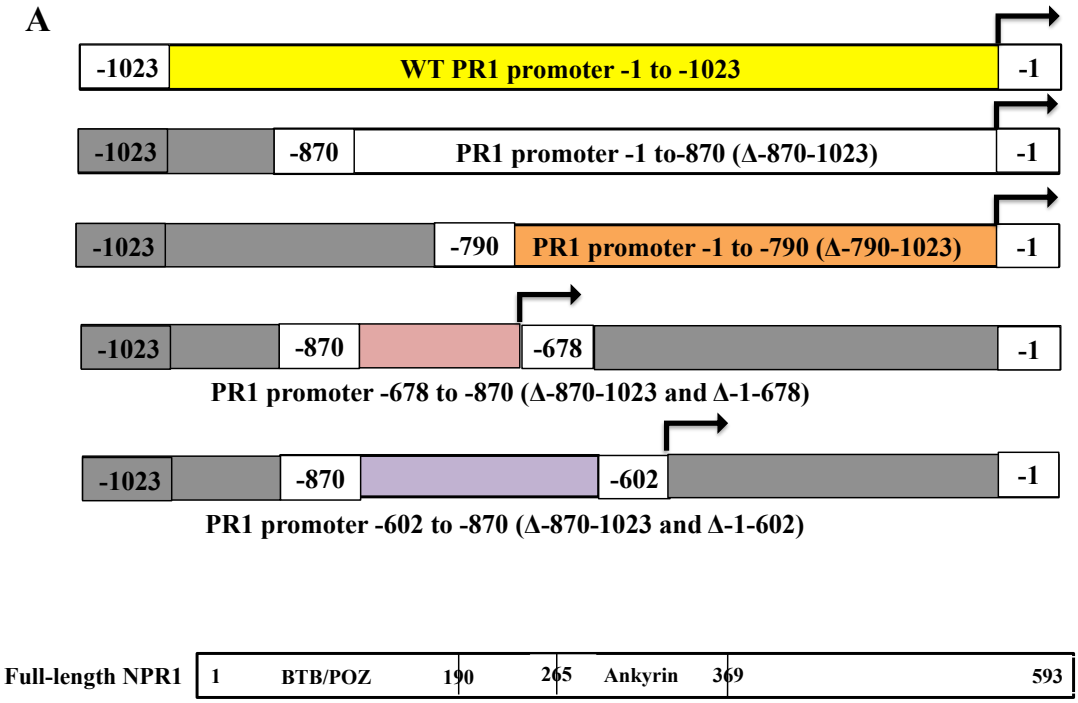
APPENDIX

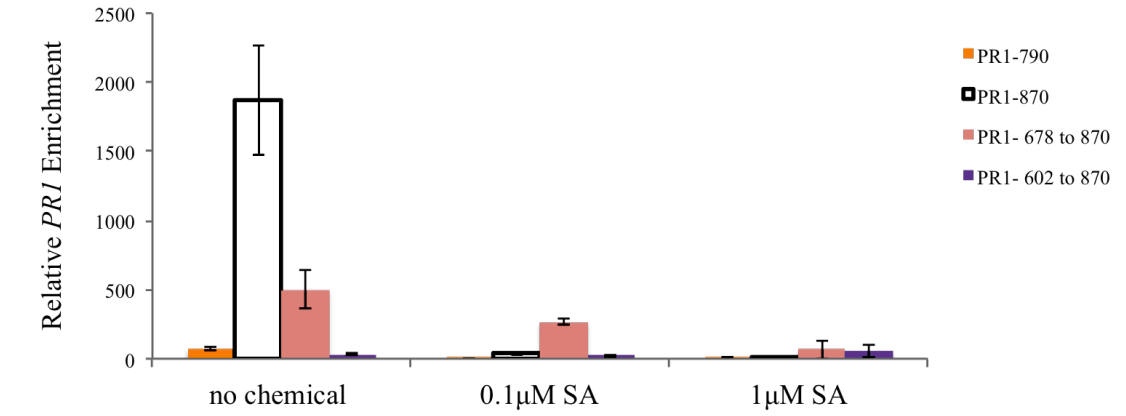


Supplemental Figure 1. Interaction of NPR1 with PR1 promoter in wild type *Arabidopsis* plants in presence and absence of SA and independent of TGA2/5/6

A. Chromatin Immunoprecipitation (ChIP) of NPR1 performed prior to SA treatment and after 6 hours treatment with 1mM SA on 6 weeks old wild type and *npr1-3* mutant *Arabidopsis* leaves by PCR followed by gel electrophoresis. ChIPs were performed with anti-NPR1 antibody. PI indicates that ChIP was performed with pre-immune serum. PCR was conducted with PR1 promoter-specific primers. The arrow indicates the location of the PCR products. The NPRl-3 protein is a deletion version of NPR1 (Cao et al., 1997), which has lost the antigenic region used to raise the anti-NPR1 antibodies used in this study. The inputs represent 2% of the immunoprecipitated material (50-fold dilution).
B. Chromatin Immunoprecipitation (ChIP) of NPR1 performed prior to SA treatment and after 6 hours treatment with 1mM SA on 6 weeks old *tga2/5/6*knockout *Arabidopsis* leaves by PCR followed by gel electrophoresis. ChIPs were performed with anti- NPR1 antibody. PI indicates that ChIP was performed with pre-immune serum. PCR was conducted with PR1 promoter-specific primers. The arrow indicates the location of the PCR products. The inputs represent 2% of the immunoprecipitated material (50-fold dilution). 5X indicate that the PCR reaction was performed with 5 times the amount of immunoprecipitated material, to demonstrate that the PCR reaction was in the linear range.

(The figure is from Rochon et al., (2006) The Plant Cell1**8**: 3670-3685).





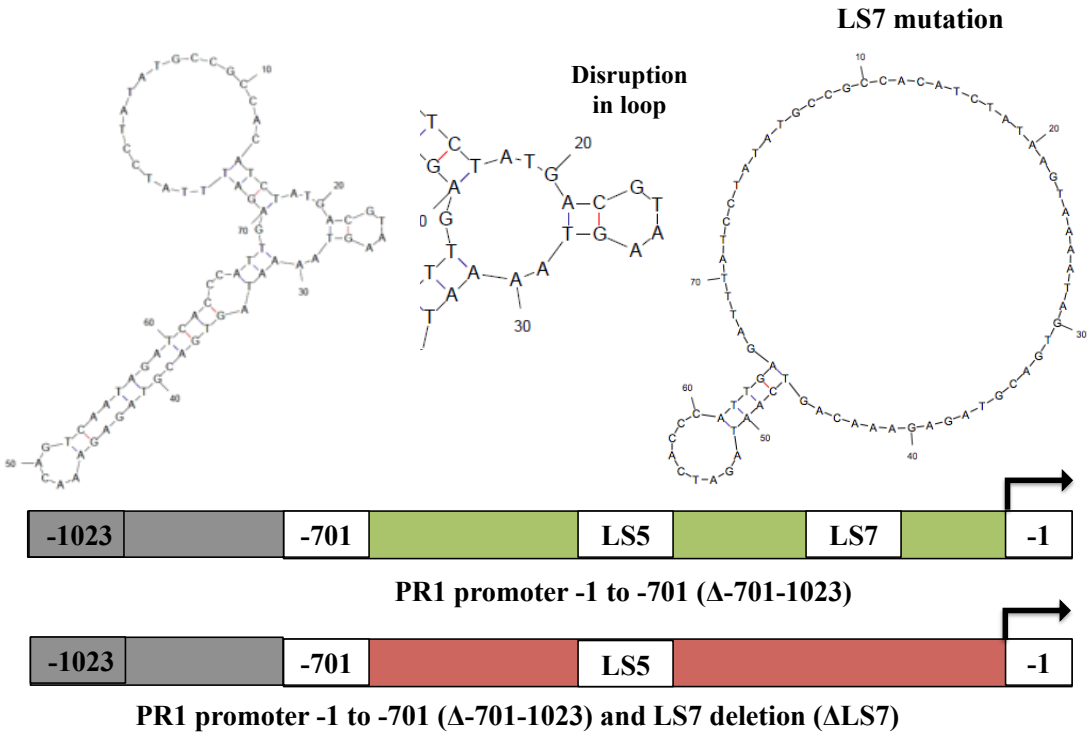
Supplemental Figure 2. NPR1 binds to -678 to -870 fragment of PR1 promoter in presence of SA

A. Schematic representation of PR1 promoter deletions constructs used for modified in vitro ChIP experiments. The arrow indicates the start of transcription. Each box represents a fragment of PR1 promoter sequence cloned into a plasmid and used as DNA source in modified ChIPs. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). White box indicates the -1 to -870bp fragment of the PR1 promoter, which sequences from -870bp to -1023bp deleted (Grey colour box). Orange box indicates the -1 to -790bp fragment of the PR1 promoter, which sequences from -790bp to -1023bp deleted (Grey colour box). Red box indicates

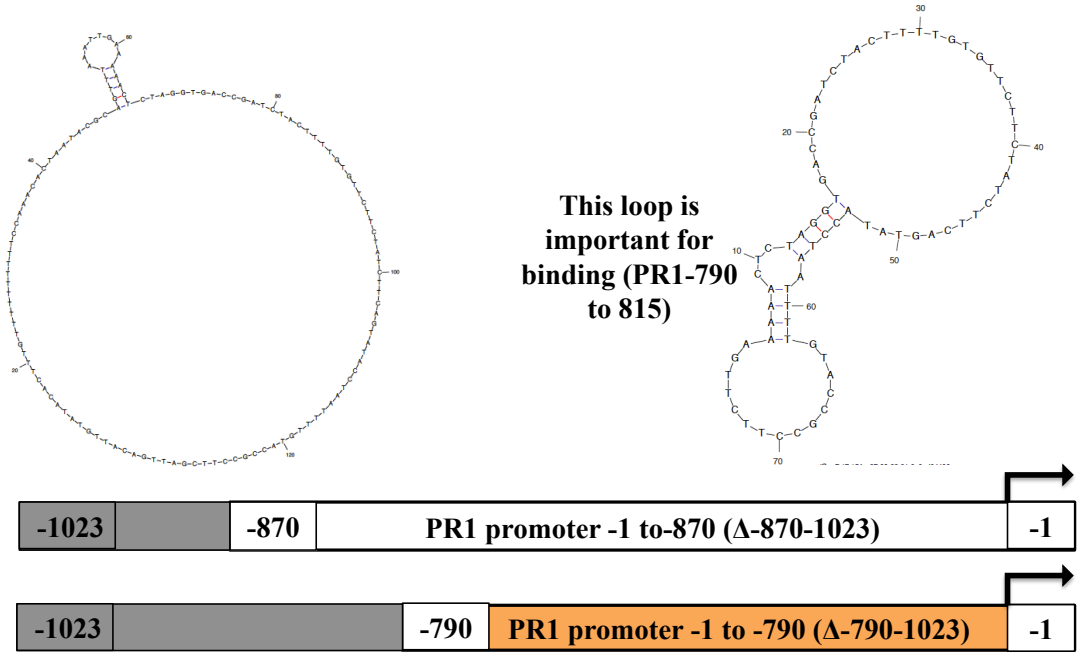
the -678bp to -870bp fragment of the PR1 promoter, which sequences from -1bp to -678bp and -870bp to -1023bp deleted (Grey colour box). Purple box indicates the -602bp to -870bp fragment of the PR1 promoter that sequences from -1bp to -602bp and -870bp to -1023bp deleted (Grey colour box). The last white box shows full-length NPR1 protein (amino acids 1-593).

B. Histogram illustrating modified ChIPs performed on PR1 promoter deletions constructs confirming NPR1 binding to -678 to -870 fragment of PR1 promoter in presence of SA. ChIPs were performed with anti-Strep antibody. Error bars indicate mean ± 1SD. Values represent averages of 3 biological replicates.

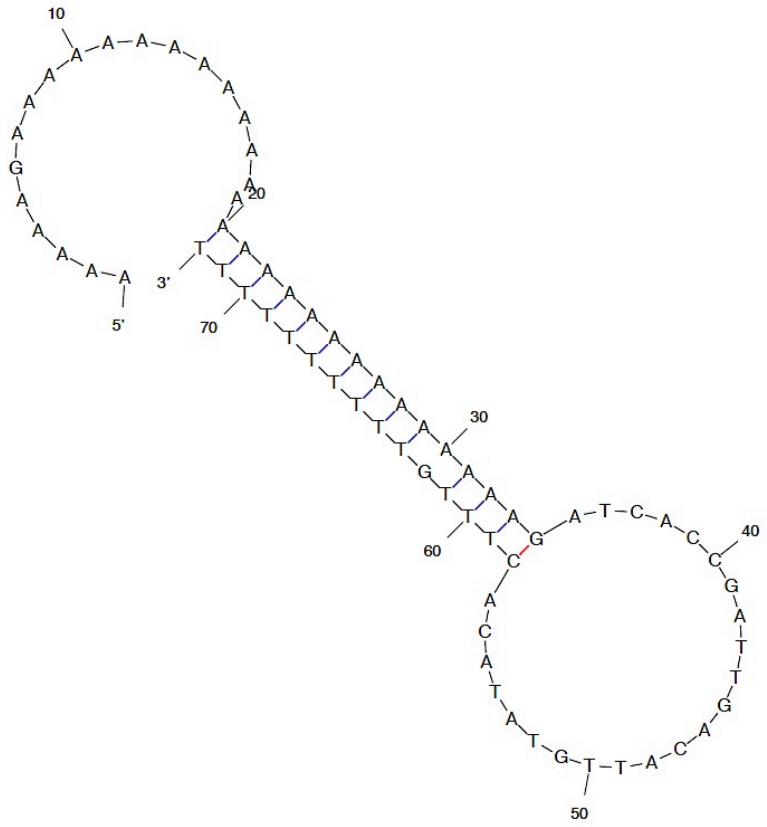
A


LS7 mutation
Disruption in loop
-1023
-701
LS5
LS7
-1
PR1 promoter -1 to -701 (Δ-701-1023)
-1023
-701
LS5
-1
PR1 promoter -1 to -701 (Δ-701-1023) and LS7 deletion (ΔLS7)

B


This loop is important for binding (PR1-790 to 815)
-1023
-870
PR1 promoter -1 to-870 (Δ-870-1023)
-1
-1023
-790
PR1 promoter -1 to -790 (Δ-790-1023)
-1

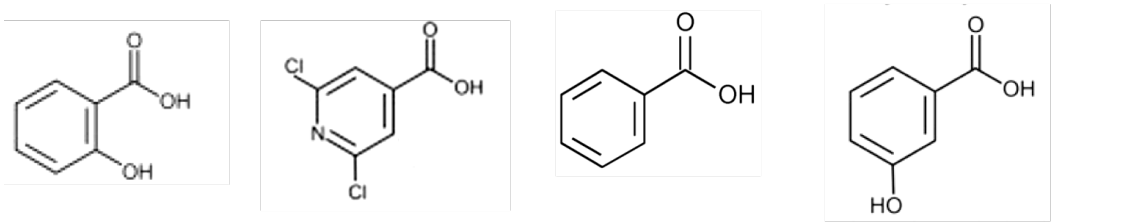
C



Supplemental Figure 3. Predicted secondary structures formation around NPR1 binding sites to PR1 promoter

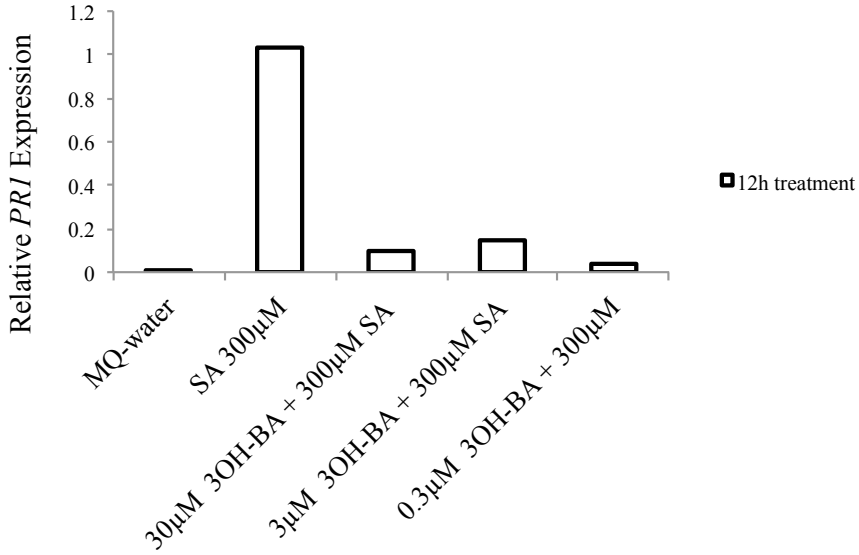
A. Multi loop formation around -701 fragment of PR1 promoter. Mfold DNA secondary structure prediction software predicted formation of stem-loop around LS7 site, where NPR1 binds to -701 fragment of PR1 promoter in presence of SA. Mutation is LS7 site (TGACG) resulted in disruption of small loops formation and loss of NPR1 binding.
B. Multi loop formation around -870 fragment of PR1 promoter. Mfold DNA secondary structure prediction software predicted formation of stem-loop at -790 to -815 fragment, where NPR1 protein binds to -870 fragment of PR1 promoter in absence of SA. Disruptions in small loop at -790 fragment resulted in complete loss of NPR1 binding.
C. Stem-loop (hairpin) structure formation at A/T rich repeats at -700 to -800 fragment on PR1 promoter. We used Mfold DNA secondary structure prediction software.

A

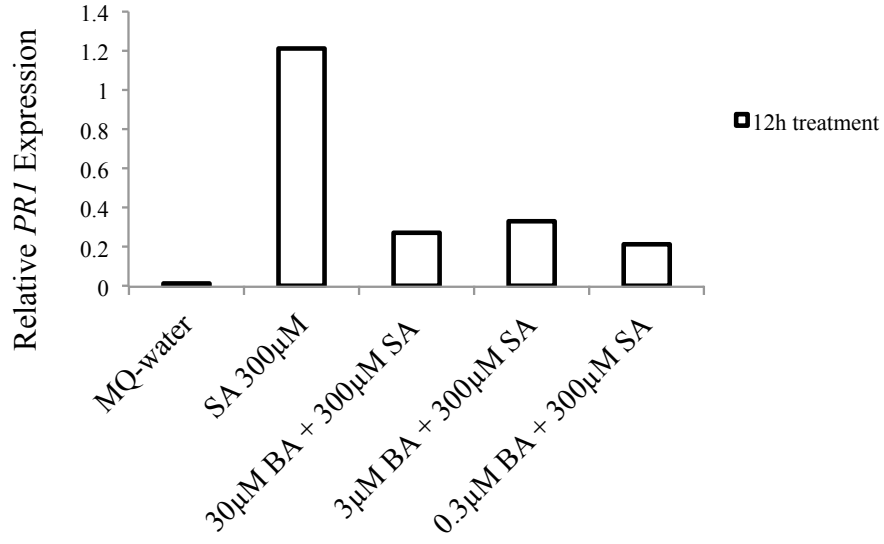


Salicylic acid (SA) | **2,6 Dichloro INA (INA)** | **Benzoic acid (BA)** | **3 Hydroxy Benzoic acid (3-OH-BA)**

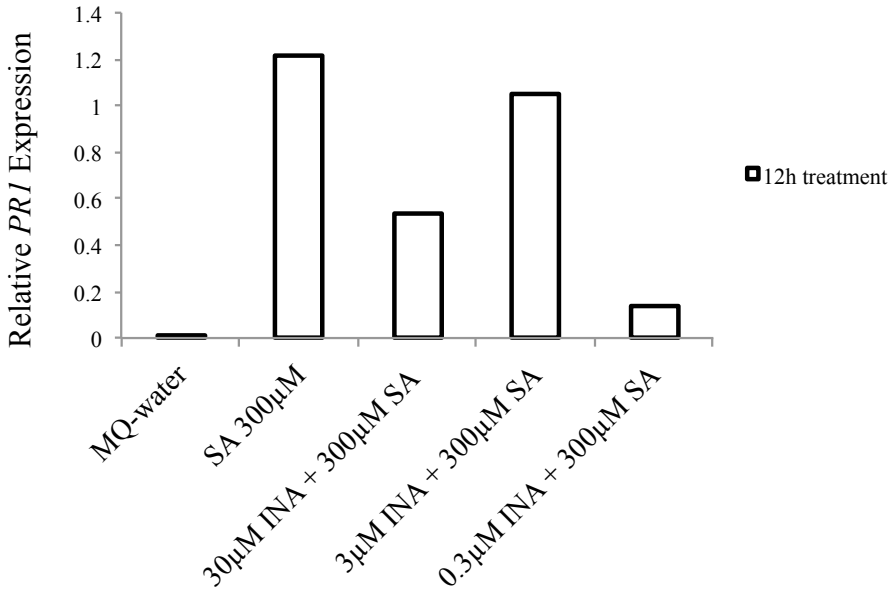
B



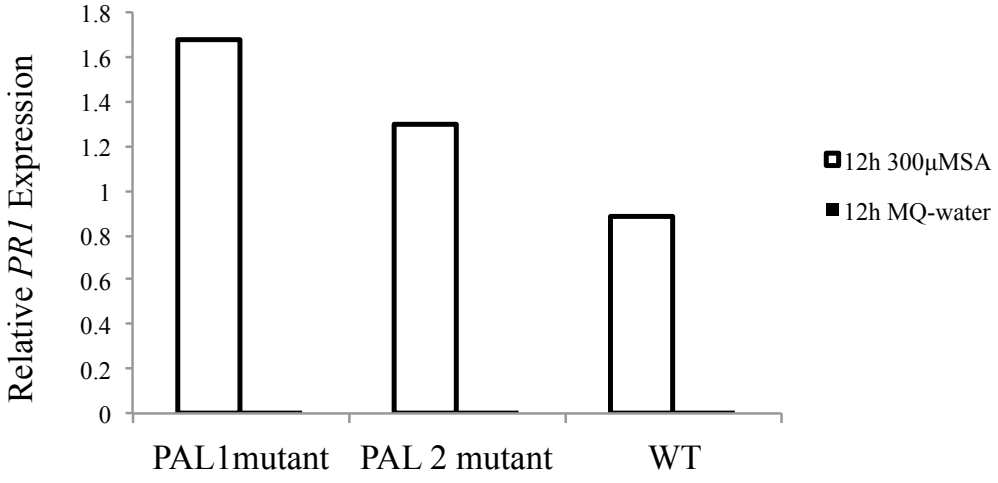
C



D



E



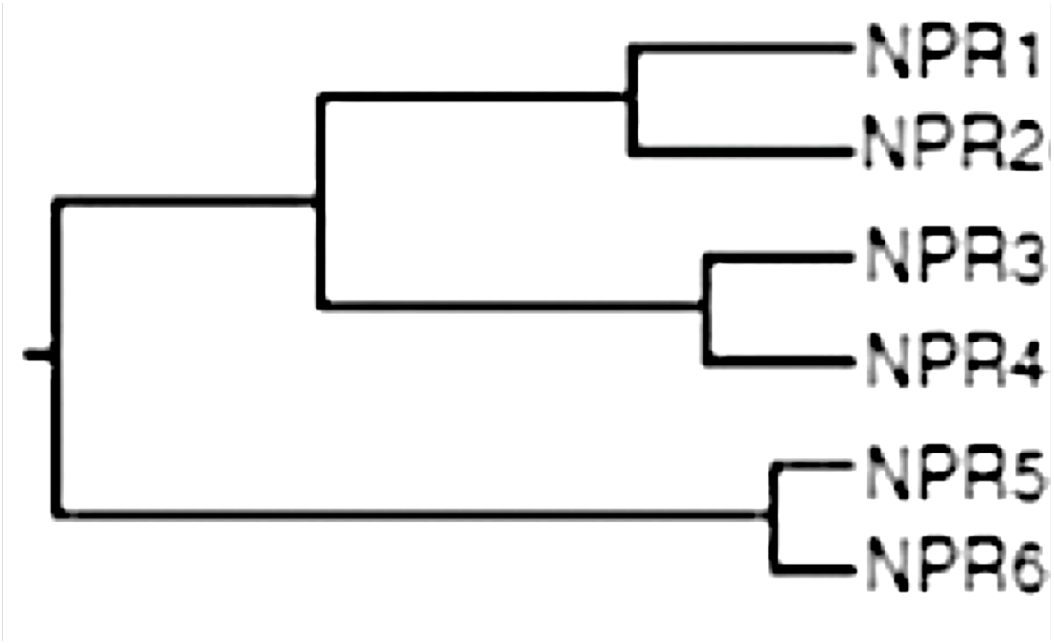
Supplemental Figure 4. *PR1* expression suppression by INA, BA, and 3OH-BA

A. Structures of SA, INA, BA, and 3OH-BA.
B. Histogram illustrating the expression profile of *PR1* gene prior to SA treatment, 12 hours after just 300µM SA treatment, and 12 hours after a combination of 300µM SA with 30, 3, and 0.3µM 3OH-BA by using the quantitative reverse transcriptase PCR (RT-PCR) technique in 6 weeks old *Arabidopsis thaliana* leaves. Expression was normalized to the expression of the constitutive gene *ubiquitin 5*.
C. Histogram illustrating the expression profile of *PR1* gene prior to SA treatment, 12 hours after just 300µM SA treatment, and 12 hours after a combination of 300µM SA with 30, 3, and 0.3µM BA by using the quantitative reverse transcriptase PCR (RT-PCR) technique in 6 weeks old *Arabidopsis thaliana* leaves. Expression was normalized to the expression of the constitutive gene *ubiquitin 5*.
D. Histogram illustrating the expression profile of *PR1* gene prior to SA treatment, 12 hours after just 300µM SA treatment, and 12 hours after a combination of 300µM SA with 30, 3, and 0.3µM INA by using the quantitative reverse transcriptase PCR (RT-PCR)

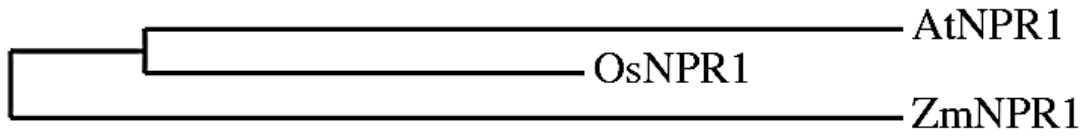
technique in 6 weeks old *Arabidopsis thaliana* leaves. Expression was normalized to the expression of the constitutive gene *ubiquitin 5*.

E. Histogram illustrating the expression profile of *PR1* gene prior to SA treatment, and 12 hours after 300μM SA treatment by using the quantitative reverse transcriptase PCR (RT-PCR) technique in 6 weeks old *Arabidopsis thaliana* leaves. Expression was normalized to the expression of the constitutive gene *ubiquitin 5*.

A



B



Supplemental Figure 5. Phylogenetic trees of NPR1-like gene in *Arabidopsis thaliana*, *Zea mays* (Corn), and *Oryza sativa* (Rice)

A. Phylogenetic tree of six NPR1-like gene in *Arabidopsis*. The tree is made using the ClustalW program based on the amino acid sequence alignments of NPR1, 2, 3, 4, 5, and 6 proteins. (Adopted from Liu et al., 2005 with modification).
B. Phylogenetic tree of *Arabidopsis thaliana*, *Zea mays* (Corn), and *Oryza sativa* (Rice). The tree is made based on the amino acid sequence alignments of NPR1 from rice, maize, and *Arabidopsis* using Phylogeny.fr tool (Dereeper et al., 2008).

```
NPR1            ----MDTTIDGFADSYEISSTSFVATDNTDSSIVYLAAEQVLTGPDVSALQLLSNSFESV 56
NPR2            MATTTTTTTARFSDSYEFSNTSGNSFFAAESSLDYP--TEFLTPPEVSALKLLSNCLESV 58
NPR3            -----MATLTEPSSSLSFTSSHFSYGSIGSNHFSSS----SASNPEVVSLTKLSSNLEQL 51
NPR4            ----MAATAIEPSSSISFTSSHLSNPSPVVTTYHS------AAN-----LEELSSNLEQL 45
NPR5            ----------------------------------------MSN-LEESLRSLSLDFLNL 18
NPR6            ----------------------------------------MSNTFEESLKSMSLDYLNL 19
:      *  :*    .:

NPR1            FDSPDDFYSDAKLVLSD-GREVSFHRCVLSARSSFFKSALA------------AAKKEKD 103
NPR2            FDSPETFYSDAKLVLAG-GREVSFHRCILSARIPVFKSALA------------TVKEQKS 105
NPR3            LSNSDCDYSDAEIIVDG--VPVGVHRCILAARSKFFQDLFK------------KEKKISK 97
NPR4            LTNPDCDYTDAEIIIEEEANPVSVHRCVLAARSKFFLDLFK------------KDKDSSE 93
NPR5            LIN-GQAFSDVTFSVEG--RLVHAHRCILAARSLFFRKFFCGTDSPQPVTGIDPTQHGSV 75
NPR6            LIN-GQAFSDVTFSVEG--RLVHAHRCILAARSLFFRKFFCESDPSQP--GAEPANQ--- 71
: .    ::*. : :      *  ***:*:**  .* . :               :.

NPR1            SNNTAAVKLELKEIAKDYEVGFDSVVTVLAYVYSSRVRPPPKGVS---ECADENCCHVAC 160
NPR2            S---TTVKLQLKEIARDYEVGFDSVVAVLAYVYSGRVRSPPKGAS---ACVDDDCCHVAC 159
NPR3            TE---KPKYQLREMLPYGAVAHEAFLYFLSYIYTGRLKPFPLEVS---TCVDPVCSHDCC 151
NPR4            K----KPKYQMKDLLPYGNVGREAFLHFLSYIYTGRLKPFPIEVS---TCVDSVCAHDSC 146
NPR5            PASPTRGSTAPAGIIPVNSVGYEVFLLLLQFLYSGQVSIVPQKHEPRPNCGERGCWHTHC 135
NPR6            TGSGAR-AAAVGGVIPVNSVGYEVFLLLLQFLYSGQVSIVPHKHEPRSNCGDRGCWHTHC 130
:     *. : .: .* ::*:.::   *   .    * :  * *  *

NPR1            RPAVDFMLEVLYLAFIFKIPELITLYQRHLLDVVDKVVIEDTLVILKLANICGKACMKLL 220
NPR2            RSKVDFMVEVLYLSFVFQIQELVTLYERQFLEIVDKVVVEDILVIFKLDTLCGTTYKKLL 219
NPR3            RPAIDFVVQLMYASSVLQVPELVSSFQRRLCNFVEKTLVENVLPILMVAFNC--KLTQLL 209
NPR4            KPAIDFAVELMYASFVFQIPDLVSSFQRKLRNYVEKSLVENVLPILLVAFHC--DLTQLL 204
NPR5            SAAVDLALDTLAASRYFGVEQLALLTQKQLASMVEKASIEDVMKVLIASRKQ--DMHQLW 193
NPR6            TAAVDLSLDILAAARYFGVEQLALLTQKHLTSMVEKASIEDVMKVLIASRKQ--DMHQLW 188
. :*: :: :  :  : : :*    :::: . *:*  :*: : ::           :*

NPR1            DRCKEIIVKSNVDMVSLEKSLPEELVKEIIDRRKELGLE------------VPKVKK--- 265
NPR2            DRCIEIIVKSDIELVSLEKSLPQHIFKQIIDIREALCLE------------PPKLER--- 264
NPR3            DQCIERVARSDLYRFCIEKEVPPEVAEKIKQLRLISPQDE---------ETSPKISEKLL 260
NPR4            DQCIERVARSDLDRFCIEKELPLEVLEKIKQLRVKSVN-------------IPEVEDKSI 251
NPR5            TTCSHLVAKSGLPPEILAKHLPIDVVTKIEELRLKSSIARRSLMPHNHHHDLSVAQDLED 253
NPR6            TTCSYLIAKSGLPQEILAKHLPIELVAKIEELRLKSSMPLRSLMPH--HHDLTSTLDLED 246
*   :.:*.:    : * :* .:  :* : *                    .

NPR1            -HVSNVHKALDSDDIELVKLLLKEDHTNLDDACALHFAVAYCNVKTATDLLKLDLADVNH 324
NPR2            -HVKNIYKALDSDDVELVKMLLLEGHTNLDEAYALHFAIAHCAVKTAYDLLELELADVNL 323
NPR3            ERIGKILKALDSDDVELVKLLLTESDITLDQANGLHYSVVYSDPKVVAEILALDMGDVNY 320
NPR4            ERTGKVLKALDSDDVELVKLLLTESDITLDQANGLHYAVAYSDPKVVTQVLDLDMADVNF 311
NPR5            QKIRRMRRALDSSDVELVKLMVMGEGLNLDESLALHYAVESCSREVVKALLELGAADVNY 313
NPR6            QKIRRMRRALDSSDVELVKLMVMGEGLNLDESLALIYAVENCSREVVKALLELGAADVNY 306
:  .: :****.*:****:::     .**:: .* :::  .  :..  :* *  .***

NPR1            RN-PRGYTVLHVAAMRKEPQLILSLLEKGASASEATLEGRTALMIAKQATMAVECNNIPE 383
NPR2            RN-PRGYTVLHVAAMRKEPKLIISLLMKGANILDTTLDGRTALVIVKRLTKADDYKTSTE 382
NPR3            RN-SRGYTVLHFAAMRREPSIIISLIDKGANASEFTSDGRSAVNILRRLTNPKDYHTKTA 379
NPR4            RN-SRGYTVLHIAAMRREPTIIIPLIQKGANASDFTFDGRSAVNICRRLTRPKDYHTKTS 370
NPR5            PAGPAGKTPLHIAAEMVSPDMVAVLLDHHADPNVRTVGGITPLDILRTLTS--------- 364
NPR6            PAGPTGKTALHIAAEMVSPDMVAVLLDHHADPNVQTVDGITPLDILRTLTS--------- 357
. * * **.**   .* ::  *: : *.    *  * :.: * :  *

NPR1            QCKHSLKGRLCVEILEQEDKREQI-PRDVPPSFAVAADELKMTLLDLENRVALAQRLFPT 442
NPR2            DGTPSLKGGLCIEVLEHEQKLEYLSPIEASLSLPVTPEELRMRLLYYENRVALARLLFPV 442
NPR3            KGRESSKARLCIDILEREIRKNPMVL-DTPMCSISMPEDLQMRLLYLEKRVGLAQLFFPT 438
NPR4            R-KEPSKYRLCIDILEREIRRNPLVSGDTPTCSHSMPEDLQMRLLYLEKRVGLAQLFFPA 429
NPR5            ----------------------DFLFKGAVPGLTHIEPNKLRLCLELVQS---------- 392
NPR6            ----------------------DFLFKGAIPGLTHIEPNKLRLCLELVQS---------- 385
: :     .      .:.*:: *   :.

NPR1            EAQAAMEIAEMKGTCEFIVTS-LEPDRLTGTKRTSPGVKIAPFRILEEHQSRLKALSKTV 501
NPR2            ETETVQGIAKLEETCEFTASS-LEPDHHIGEKRTSLDLNMAPFQIHEKHLSRLRALCKTV 501
NPR3            EAKVAMDIGNVEGTSEFTGLS--PPSSGLTGNLSQVDLNETPHMQTQRLLTRMVALMKTV 496
NPR4            EANVAMDVANVEGTSECTGLLTPPPSNDTTENLGKVDLNETPYVQTKRMLTRMKALMKTV 489
NPR5            ---AAMVISREEG-------------NNSNNQNNDNNTGIYPHMNEEHNSGSSGGSNNNL 436
NPR6            ---AALVISREEG-------------NNNS---NDNNTMIYPRMKDEHTSGSS------L 420
..  :.. :                     . .    *    :.          :

NPR1            ELGKRFFPRCSAVLDQIMN---CEDLTQLACGEDDTAEKRLQKKQRYMEIQETLKKAFSE 558
NPR2            ELGKRYFKRCS--LDHFMD---TEDLNHLASVEEDTPEKRLQKKQRYMELQETLMKTFSE 556
NPR3            ETGRRFFPYGSEVLDKYMAEYIDDDILDDFHFEKGSTHERRLKRMRYRELKDDVQKAYSK 556
NPR4            ETGRRYFPSCYEVLDKYMDQYMDEEIPDMSYPEKGTVKERRQKRMRYNELKNDVKKAYSK 549
NPR5            DSRLVYLNLGAGTGQMGPG----RDQGDDHNSQREGMSRHHHHHQDPSTMYHHHHQHHF- 491
```

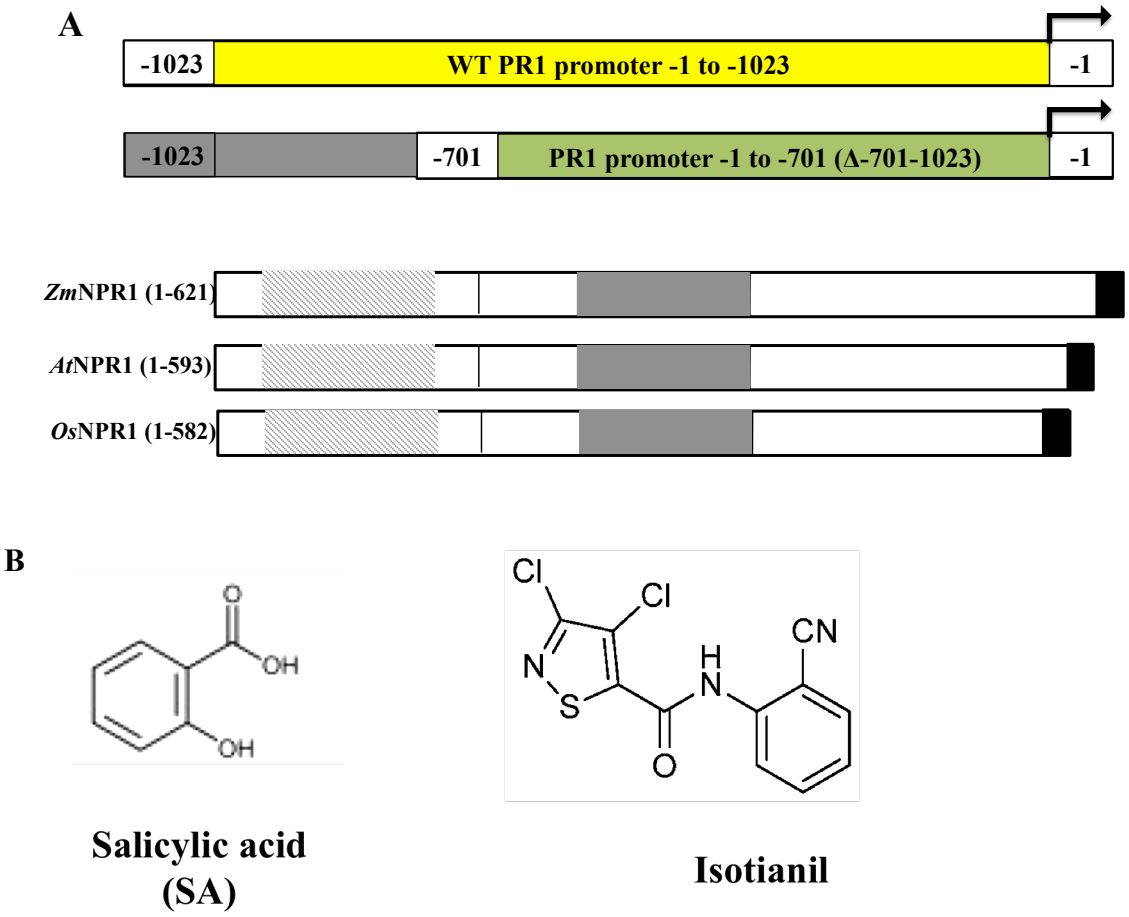
```
NPR6            DSRLVYLNLGATN----------RDIGDDNSNQREGMNLHHHHH-DPSTMYHHHHHHF-- 467
:    ::                 :  .    :      :  ::     : .   : .

NPR1            DNL-ELGNSSLTDSTSSTSKSTGGKRSNRKLSHRRR---------------- 593
NPR2            DKE-ECGKS-------STPKPTSAVRSNRKLSHRRLKVDKRDFLKRPYGNGD 600
NPR3            DKESKIARSCL-----SASSSPSSSSIRDDLHNTT----------------- 586
NPR4            DK---VARSCL-----SSSS--PASSLREALENPT----------------- 574
NPR5            ----------------------------------------------------
NPR6            ----------------------------------------------------
```

Supplemental Figure 6. Amino acid sequence alignment of all 6 *Arabidopsis thaliana* NPR proteins

The yellow highlight indicates the position corresponding to cysteines found in NPR1 essential for binding to SA. The (*) below the alignment sequences indicates identical residues, (:) indicates residues with strongly similar properties and (.) indicates residues with weakly similar properties. Alignment was created using the ClustalW2 program using sequences from UniProtKB database.

Supplemental Figure 7. Amino acid sequence alignment of *Arabidopsis thaliana* NPR1, *Zea mays* (Corn) NPR1, and *Oryza sativa* (Rice) NPR1

The yellow highlight indicates the position corresponding to cysteines found in NPR1 essential for binding to SA. The (*) below the alignment sequences indicates identical residues, (:) indicates residues with strongly similar properties and (.) indicates residues with weakly similar properties. Alignment was created using the ClustalW2 program using sequences from UniProtKB database.



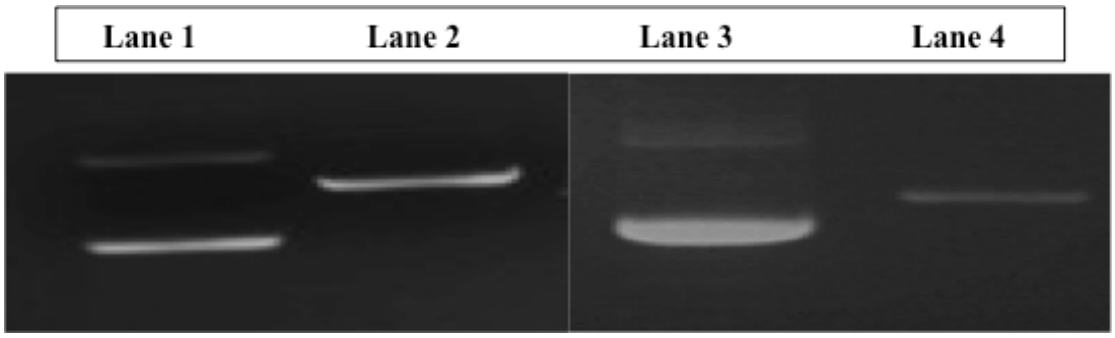
Supplemental Figure 8. *Arabidopsis*, maize and rice NPR1 proteins bind to -701 fragment of PR1 promoter in presence of SA

A. Schematic representation of PR1 promoter deletions constructs and NPR1 proteins from rice, maize, and *Arabidopsis* used for modified ChIP experiments. The yellow box represents the -1 to -1023bp fragment of the PR1 promoter sequence (referred as wild type PR1 promoter). Green box represents the -1 to -701bp fragment of the PR1 promoter that sequences from -701bp to -1023bp deleted (Grey colour box). White boxes show

Arabidopsis thaliana NPR1 protein (*At*NPR1) (amino acid 1-593), *Zea mays* NPR1 protein (*Zm*NPR1) (amino acid 1-621), and *Oryza sativa* NPR1 protein (*Os*NPR1) (amino acid 1-582). Regions with homology to known functional domain are shown in different color box. Hatched boxes represent BTB/POZ domains, grey boxes show ankyrin repeat regions, and Black boxes indicate nuclear localization sequence.

B. Structures of SA and 3,4-Dichlor-2′-cyano-1,2-thiazol-5-carboxanilid (Isotianil).

C. Histogram illustrating modified ChIP experiments performed on *Arabidopsis*, rice and maize NPR1 proteins for binding to -701 fragment of PR1 promoter before and after addition of SA and Isotianil, confirming that all three NPR1 proteins bind to -701 fragment only in presence of SA. ChIPs were performed with anti-Strep antibody. The relative *PR1* enrichment analysed by qPCR.



Supplemental Figure 9. Confirmation of -701 and -870 fragments of PR1 promoter before and after digestion with *Bgl*I restriction enzyme

Digestion with *Bgl*I linearized the vector and produced a single band on gel. Lane 1 shows -701-fragment before digestion. Lane 2 indicates -701-fragment after digestion with *Bgl*I. Lane 3 shows -870-fragment of PR1 promoter before digestion. Lane 4 represents -870-fragment after digestion with *Bgl*I.

Supplemental Table 1. Sequence identity percentages of six members of the NPR family gene in *Arabidopsis thaliana* generated using ClustalW2 multiple sequence alignment tool.

NPR1	NPR2	NPR3	NPR4	NPR5	NPR6
NPR1	61.89	35.67	36.24	20.16	21.63
NPR2		36.52	37.80	18.94	20.34
NPR3			71.95	24.24	25.48
NPR4				24.24	25.91
NPR5					83.51
NPR6					

Supplemental Table 2. Sequence identity percentages of NPR1-like gene in *Arabidopsis thaliana* (*At*NPR1), *Zea mays* (Corn)(*Zm*NPR1), and *Oryza sativa* (Rice)(*Os*NPR1) generated using ClustalW2 multiple sequence alignment tool.

	***At*NPR1**	***Zm*NPR1**
***Os*NPR1**	45.88	40.72
***At*NPR1**		35.75
***Zm*NPR1**		

Supplemental Table 3. Sequence of the primers used for constructing the PR1 promoter deletions constructs and qPCR analysis in modified *in vitro* ChIP experiments.

Primer Name	Sequence
PR1T-8a	GGACAGTTTGGCAATTAAGA
PR1T-8b	GCATGAACACTAAGAAACTT
PR1T-9a	CTTAGTGTTTCATGCATATGA
PR1T-9b	GACTGTTTCTCTACGTCACTA
PR1T-10a	GCATATGAGTATCTCTATCACTC
PR1T-10b	GGTGATCTATTGACTGTTTCT
PR1T-11a	CACTCTTGCCTATGGCTGAA
PR1T-11b	TCAATTTAAACTGCGTATTA
PR1T-12a	CGCCACATCTATGACGTAAG
PR1T-12b	TACAATGTCAATCGGTGATC
PR1T-13a	TAGTGACGTAGAGAAACAGTC
PR1T-13b	GTGTATACAATGTCAATCGGT
PR1T-14a	AGATCACCCATTGAGATTTAT
PR1T-14b	ACTGCGTATTAGTGTTTGGAA
PR1T-15a	GATCACCGATTGACATTGTA
PR1T-15b	GAACACAAAAGTAGATCGGT
PR1T-16a	AAAAGGTTTGTGATTATGCG
PR1T-16b	AAGGCGGTACAAAATTAGGT
UBQ5a	GACGCTTCATCTCGTCC
UBQ5b	GTAAACGTAGGTGAGTCCA